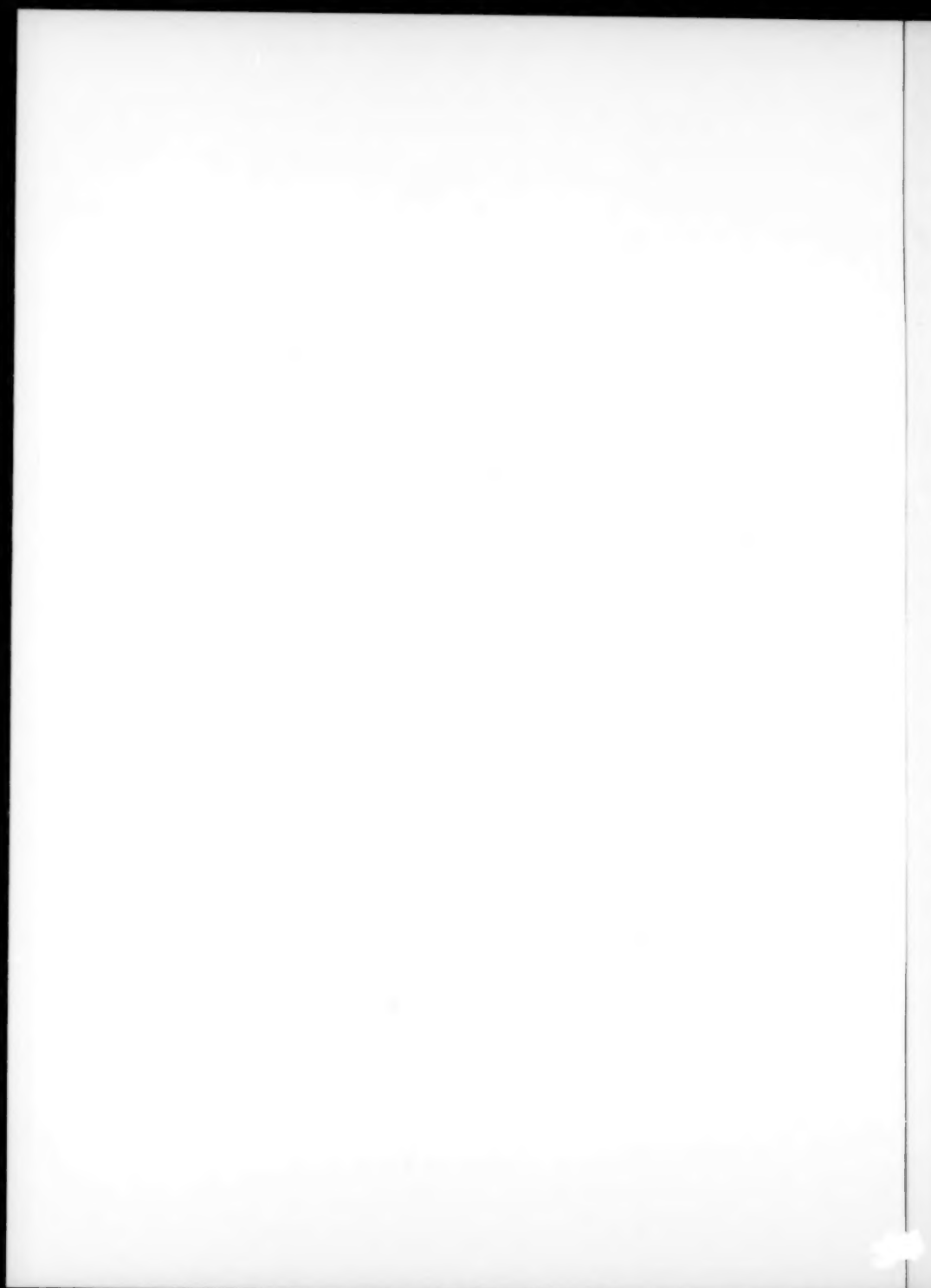


**MATHEMATISCHE
ANNALEN**

118. BAND



MATHEMATISCHE ANNALEN

BEGRÜNDET 1868 DURCH
ALFRED CLEBSCH UND CARL NEUMANN

FORTGEFÜHRT DURCH
FELIX KLEIN DAVID HILBERT

GEGENWÄRTIG VERTRETUNGSWEISE
HERAUSGEGEBEN VON

ERICH HECKE
IN HAMBURG

UNTER MITWIRKUNG VON
HEINRICH BEHNKE BARTEL L. VAN DER WAERDEN
IN MÜNSTER I. W. IN LEIPZIG

118. BAND



BERLIN
VERLAG VON JULIUS SPRINGER
1941/1943

Unveränderter Nachdruck 1975

Springer-Verlag Berlin Heidelberg New York



Inhalt des 118. Bandes.

(In alphabetischer Ordnung.)

	Seite
Alexits, G. v., Über verstreute Mengen	379
(Anschrift: Budapest XII, Györi-ut 12)	
Alfred Ackermann-Teubner-Gedächtnispreis	440
Behrens, E.-A., Über die Existenz von Algebren beliebigen Ranges mit quadratischer Normenform	85
(Anschrift: Görlitz, Eichendorffstr. 2)	
Borůvka, O., Über Ketten von Faktoroiden	41
(Anschrift: Brünn XVI, Matzenauergasse 3)	
Bullig, G., Zur Kettenbruchtheorie im n -Dimensionalen (Z 3)	1
(Anschrift: Hamburg, Rothenbaumchaussée 21, Mathematisches Seminar)	
Dinghas, A., Zum isoperimetrischen Problem für die nichteuklidischen Geometrien	636
(Anschrift: Berlin C 2, Mathem. Institut der Universität, Unter den Linden 6)	
Dörge, K., Beweis des Reziprozitätsgesetzes für quadratische Reste	310
(Anschrift: Bensberg b. Köln, Kölner Str. 95)	
Hadwiger, H., Über die Umordnungsstärke und eine Erweiterung des Steinitschen Satzes	702
(Anschrift: Bern, Schweiz, Hochfeldstr. 31)	
Haupt, O., Über den Ovalsatz von Böhmer-Mukhopadhyaya	629
(Anschrift: Erlangen, Spardorfer Str. 45)	
Iglisch, R., Bemerkungen zu einigen von Herrn Collatz angegebenen Eigenwertabschätzungen bei linearen Integralgleichungen	263
(Anschrift: Braunschweig, Wilhelm-Bode-Str. 12)	
Keller, O.-H., Eine Bemerkung zu den Plückerschen Formeln	626
(Anschrift: Mürwik, Marineschule)	
Kerékjártó, B. von, Über die dreigliedrigen integrierbaren Gruppen	365
(Anschrift: Budapest VIII, Muzeum-körül 6—8)	
Kloosterman, H. D., Simultane Darstellung zweier ganzen Zahlen als einer Summe von ganzen Zahlen und deren Quadratsumme	319
(Anschrift: Leiden, Holl., van Oldenbarneveltstraat 52)	
Kurepa, G., Über eine Eigenschaft von Systemen linearer wohlgeordneter Mengen	578
[Anschrift: Zagreb (Agram), Kroatien, Philosophische Fakultät]	
Lamuel, E., Über Approximation meromorpher Funktionen durch rationale Funktionen	134
(Anschrift: Prag I, Dominikanergasse 5, Deutsche Technische Hochschule)	
Li, E. P., Die 28 Doppeltangenten einer Kurve vierter Ordnung	94
(Anschrift: Leipzig S 3, Scharnhorststr. 40 ^{II})	
Lipka, St., Über die Irreduzibilität von Polynomen	235
(Anschrift: Szeged, Ungarn, Baross-Str. 2, Universität)	
Lipka, St., Integralsätze über Polynome mit lauter reellen Nullstellen	485
(Anschrift: Szeged, Ungarn, Baross-Str. 2, Universität)	
Maak, W., Schnittpunktzahl rektifizierbarer und nichtrektifizierbarer Kurven	299
(Anschrift: Hamburg 13, Mathemat. Sem. d. Univ.)	
Maas, H., Modulformen und quadratische Formen über dem quadratischen Zahlkörper $R(\sqrt{5})$	65
(Anschrift: Heidelberg, Hauptstr. 76 ^{II})	
Maas, H., Über eine Metrik im Siegelschen Halbraum	312
(Anschrift: Heidelberg, Schillerstr. 37 ^{II})	
Maas, H., Theorie der Poincaréschen Reihen zu den hyperbolischen Fixpunktsystemen der Hilbertschen Modulgruppe	518
(Anschrift: Heidelberg, Schillerstr. 37 ^{II})	
Nakano, H., Unitärinvarianten im allgemeinen Euklidischen Raum	112
(Anschrift: Tokyo, Japan, Mathematisches inst. Faculty of Science Tokyo Imperial University)	

	Seite
Neder, L., Modell einer Differentialrechnung mit aktual unendlich kleinen Größen erster Ordnung	251
(Anschrift: Münster i. W., Gertrudenstr. 39)	
Neder, L., Modell einer Leibnizischen Differentialrechnung mit aktual unendlich kleinen Größen sämtlicher Ordnungen	718
(Anschrift: Münster i. W., Gertrudenstr. 39)	
Neumann, E. R., Inversion und konforme Abbildung von Komplementärgebieten	276
(Anschrift: Marburg a. d. Lahn, Schwanenallee 9)	
Nöbeling, G., Über die Flächenmaße im Euklidischen Raum	687
(Anschrift: Erlangen, Lammerstr. 5)	
Novák, J., Induktion partiell stetiger Funktionen	449
(Anschrift: Brünn, Parkstr. 56)	
Perron, O., Über die Bedingungen, daß eine binäre Form n -ten Grades eine n -te Potenz ist, und über die rationale Kurve n -ter Ordnung im R_n	306
(Anschrift: München 27, Friedrich-Herschel-Str. 11)	
Perron, O., Beweis und Verschärfung eines Satzes von Kronecker	441
(Anschrift: München 27, Friedrich-Herschel-Str. 11)	
Pospíšil, B., Von den Verteilungen auf Booleschen Ringen	32
(Anschrift: Brünn, Sirotčí 36)	
Preisauflage der Technischen Hochschule Karlsruhe	608
Rédei, L., Zur Frage des Euklidischen Algorithmus in quadratischen Zahlkörpern	588
(Anschrift: Szeged, Ungarn, Baross Gábor u. 2 Sz.)	
Rellich, F., Störungstheorie der Spektralzerlegung. V.	462
(Anschrift: Dresden, Münchener Str. 8 b)	
Richter, H., Untersuchungen zum Erneuerungsproblem	145
(Anschrift: Leipzig S 3, Kronprinzenstr. 106)	
Sauer, L., Parametrixmethode zur Lösung von Randwertproblemen	385
(Anschrift: Frankfurt a. M., Schumannstr. 58, Mathem. Sem. d. Univ.)	
Schmidt, A., Die Dualität von Inzidenz und Senkrechtstehen in der absoluten Geometrie	609
(Anschrift: Potsdam, Alexandrinenstr. 1 ^a)	
Schrutka, L. v., Eine neue Einteilung der Permutationen	246
(Anschrift: Wien 98, Onno-Klopp-Gasse 8, Techn. Hochschule)	
Schwarz, H., Zur „wahrscheinlichkeitstheoretischen Stabilisierung“ beim Erneuerungsproblem	771
(Anschrift: Göttingen, Schillerstr. 3 ¹¹)	
Schwarz, J. von, Eine Methode zur Verallgemeinerung der gewöhnlichen Differentialgleichungen	497
(Anschrift: Berlin-Friedenau, Kaiserallee 77 ¹¹)	
Tautz, G., Zur Theorie der elliptischen Differentialgleichungen. II.	733
(Anschrift: Breslau, Auenstr. 18)	
Tietze, H., Über die Anzahl der Lösungen gewisser Aufgaben über Gitterpunktfiguren	290
(Anschrift: München 23, Trautenwolfstr. 7 ¹¹)	
Waerden, B. L. vander, Zur pythagoreischen Algebra: Quadratwurzel und Kubikwurzel	286
(Anschrift: Leipzig S 3, Fockestr. 8 a)	
Wecken, F., Fixpunktklassen. II	216
(Anschrift: Marburg a. d. Lahn, Frankfurter Str. 38 ¹¹)	
Wecken, F., Fixpunktklassen. Teil III. Mindestzahlen von Fixpunkten	544
(Anschrift: Marburg a. d. Lahn, Frankfurter Str. 38 ¹¹)	
Wintgen, G., Zur Darstellungstheorie der Raumgruppen	195
(Anschrift: Leipzig, Mathem. Seminar der Universität)	
Zu Pölberts 80. Geburtstag	289
Todesanzeige für David Hilbert	nach 608

Zur Kettenbruchtheorie im n -Dimensionalen (Z3).

Von

G. Bullig in Hamburg.

Diese Arbeit bringt die Verallgemeinerung der in Z 1¹⁾ für drei Dimensionen behandelten Theorie auf n Dimensionen. Hierbei wird $n \geq 3$ vorausgesetzt, da für $n = 2$ nur ein geringer Bestandteil der Theorie, der sich übrigens leicht aus den folgenden Untersuchungen herauslösen läßt, Bedeutung hat.

Bezüglich einer vorgegebenen Menge M , die gewisse einschränkende Forderungen erfüllt (S. 6, 7), werden extreme Quader (n -ter Art, S. 7) definiert. Diese Quader enthalten im Innern keinen Punkt von M , auf n bestimmten ihre Seitenflächen aber je einen. Im Falle, wo M aus Punkten eines Gitters besteht, liefert also jeder extreme Quader n Gitterpunkte. Überträgt man den Begriff des extremen Quaders auf die 2. Dimension, so erhält man bei geeigneter Wahl von M in den Punktepaaren der Seiten der extremen Quader genau die reduzierten Basen des Kettenbruchverfahrens. In drei Dimensionen sind die durch die extremen Quader definierten Tripel von Gitterpunkten bei geeigneter Wahl von M ebenfalls Basispunkte des Gitters, und die Bedingung, daß ein Quader extrem ist, läßt sich wie in 2. Dimension durch Ungleichungen ausdrücken, welche die Koordinaten der drei Gitterpunkte erfüllen. Darauf beruht die Existenz eines Kettenbruchverfahrens zur Bestimmung aller extremen Quader in 3. Dimension (vgl. Z 2²⁾). Aus diesen Tatsachen erhellt die Bedeutung, die den extremen Quadern hinsichtlich einer Verallgemeinerung des Kettenbruchverfahrens zukommt.

Die Transformationen $\Omega_{r,s}$ (S. 26), die in zwei und drei Dimensionen die reduzierten Basen, also die extremen Quader, wiederum in solche überführen, lassen sich durch einfache zahlengeometrische Operationen ersetzen, bei denen lediglich diejenigen Eigenschaften der extremen Quader benutzt werden, durch welche sie nach der obigen Definition charakterisiert sind. Die Verallgemeinerung dieser zahlengeometrischen Prozesse auf n Dimensionen ist nahezu eine zwangsläufige. Das so definierte Verfahren wird im folgenden topologisch untersucht. Auf Grund dessen werden in folgenden Arbeiten Eigenschaften des Verfahrens hergeleitet, die zur Bestimmung einer Basis der Einheitengruppe eines beliebigen totalreellen algebraischen Zahlkörpers führen. Um Approximationseigenschaften in n Dimensionen zu beweisen,

¹⁾ Abhandl. aus d. Math. Seminar d. Hansischen Univ. 13 (1940).

²⁾ Festband der Math. Ges. Hamburg, 1940.

wird eine Darstellung der eben genannten zahlgeometrischen Prozesse durch rationale Transformationen förderlich sein, wie sie in Z 2 für $n = 3$ gegeben wurde. Desgleichen ist eine solche Darstellung von Wert für eine bequemere numerische Durchführung dieser Prozesse. Die Erledigung dieser Frage bleibt ebenfalls späteren Untersuchungen vorbehalten.

Wir schreiten nunmehr zu einer Übersicht der topologischen Eigenschaften des Verfahrens, die in dieser Arbeit bewiesen werden.

Die Untersuchung richtet sich auf die Ermittlung von Struktureigenschaften des Graphen Γ^1 , der als Eckpunkte die extremen Quader, als Strecken die Operationen Ω_{α} , enthält. Dazu wird die Begrenzungsfläche $B(V)$ der Vereinigungsmenge V aller extremen Quader — $B(V)$ ist mit dem $(n-1)$ -dimensionalen Euklidischen Raum homöomorph — in eine Menge K punktfremder topologischer offener 0 -, 1 -, ..., $(n-1)$ -dimensionaler Kugeln zerlegt. Jede Kugel ist durch ihren Rand eindeutig bestimmt. Die Eckpunkte unseres Graphen werden durch die 0 -dimensionalen Kugeln, seine Strecken durch die 1 -dimensionalen Kugeln in $B(V)$ realisiert. Es gilt weiter der wichtige Satz, daß der Rand einer t -dimensionalen Kugel ($0 < t \leq n-1$) in endlich viele 0 -, 1 -, ..., $(t-1)$ -dimensionale Kugeln aus K zerfällt; es folgt also: Der Rand einer (abgeschlossenen) 2 -dimensionalen Kugel aus K liefert einen 1 -dimensionalen Zykel von Γ^1 . Erweitern wir Γ^1 zu Γ^2 , indem wir die 2 -dimensionalen Kugeln aus K (abgeschlossen) hinzunehmen, so ist der Rand einer 3 -dimensionalen Kugel ein 2 -dimensionaler Zykel in Γ^2 usw. Da von jeder Dimension $\leq n-1$ unendlich viele Kugeln in K existieren, ist damit die Existenz unendlich vieler 1 -, 2 -, ..., $(n-1)$ -dimensionaler Zykeln in Γ^{n-1} nachgewiesen. Γ^{n-1} ist als Punktmenge gleich $B(V)$.

Es lassen sich nun Aussagen machen, die zu einer kombinatorischen Konstruktion aller genannten Zykeln führen, wenn allein der Graph Γ^1 und die Indexpaare der seinen Strecken zugeordneten Operationen Ω_{α} bekannt sind. Diese Voraussetzung ist bei Anwendung unserer Theorie z. B. auf algebraische Zahlkörper leicht erfüllbar, weil dann in Γ^1 ein Fundamentalbereich einer Gruppe existiert, der nur endlich viele Punkte und Strecken enthält, andererseits auf zwei Punkte P', P , die mod. dieser Gruppe äquivalent sind, die genau gleichen Operationen $\Omega_{1k(1)}, \dots, \Omega_{nk(n)}$ anwendbar sind (d. h. ist $\Omega_{mk(m)}$ auf P , $\Omega_{m'k'(m)}$ auf P' anwendbar, so gilt $k(m) = k'(m)$); also auch die Indexpaare lassen sich nach Vorgabe endlich vieler Paare übersehen (vgl. auch S. 3 unten).

Was die kombinatorische Konstruktion unserer Zyklen angeht, so gilt

1. Der 0 -dimensionale Teilkomplex γ^0 von Γ^1 , der aus den 0 -dimensionalen Kugeln von K besteht, die auf dem Rande einer vorgegebenen t -dimensionalen Kugel E^t von K liegen ($t \geq 1$), ist bereits eindeutig bestimmt, wenn irgendeiner seiner Punkte, P , gegeben ist mit samt den Ziffern l_1, \dots, l_s , für die

$\Omega_{l_1 k(l_1)}(P), \dots, \Omega_{l_t k(l_t)}(P)$ zu γ^0 gehören. Sogar E^t ist durch diese Angaben eindeutig bestimmt. Zunächst gilt stets $q = t$. Ist nun ein $P_i = \Omega_{l_i k(l_i)}(P)$ vorgegeben, so sind die Ziffern l'_m ($m = 1, \dots, t$), für die $\Omega_{l'_m k(l'_m)}(P_i)$ zu γ^0 gehören, gegeben durch

$$\left. \begin{aligned} l'_m &= l_m \quad \text{für } m \neq i \\ l'_i &= k(l_i) \end{aligned} \right\} \text{ falls } k(l_i) \notin (l_1, \dots, l_t),$$

$$l'_m = l_m \quad \text{für alle } m, \quad \text{falls } k(l_i) \in (l_1, \dots, l_t).$$

Da Γ^1 und die seinen Strecken zugeordneten Indexpaare als bekannt vorausgesetzt wurden, sind die Zahlen $k(l_1), \dots, k(l_t)$ sämtlich bekannt, also auch l'_i . Wir können auf Grund unseres Satzes immer weitere Eckpunkte von γ^0 bestimmen und es gilt, daß man durch diesen Prozeß ganz γ^0 erhält.

2. Es stellt sich heraus, daß zu jedem Eckpunkt P von Γ^1 und einem beliebigen System von t (paarweise verschiedenen) Ziffern l_1, \dots, l_t , die den Zahlen $1, \dots, n$ entnommen sind, eine t -dimensionale Kugel aus K gehört, die auf ihrem Rande P und $\Omega_{l_1 k(l_1)}(P), \dots, \Omega_{l_t k(l_t)}(P)$ enthält. Die so gewonnenen Kugeln sind bei festem P paarweise verschieden.

3. Sind P und $\Omega_{l_1 k(l_1)}(P), \dots, \Omega_{l_t k(l_t)}(P)$ im Rande der t -dimensionalen Kugel E^t ($t \geq 2$) gelegen, und ist l_{q_1}, \dots, l_{q_v} irgendein System paarweise verschiedener Ziffern aus dem System l_1, \dots, l_t , $v < t$, so ist die vermöge 2. und 1. durch $P, \Omega_{l_{q_1} k(l_{q_1})}(P), \dots, \Omega_{l_{q_v} k(l_{q_v})}(P)$ bestimmte v -dimensionale Kugel E^v ganz im Rande von E^t gelegen. Im Rande von E^t liegen keine weiteren v -dimensionalen Kugeln mit P als Randpunkt.

Nunmehr übersehen wir alle in Rede stehenden t -dimensionalen Zyklen von Γ^{n-1} kombinatorisch. Wir nehmen nach 2. irgendeinen Punkt P und t Ziffern l_1, \dots, l_t . Sei E^t die durch diese Daten bestimmte t -dimensionale Kugel E^t aus K . Wir kennen nach 1. 3. den auf ihrem Rande liegenden 1-dimensionalen Teilkomplex γ^1 von Γ^1 und die zu seinen Strecken gehörenden Indexpaare. Wir nehmen von den Ziffern l_1, \dots, l_t irgendein Paar l_{q_1}, l_{q_2} . Dann liegt die durch P, l_{q_1}, l_{q_2} bestimmte 2-dimensionale Kugel nach 3. im Rande von E^t . Ebenso verfahren wir mit den übrigen Ziffernpaaren, die in l_1, \dots, l_t enthalten sind. Weitere 2-dimensionale Kugeln, die im Rande P enthalten, können nach 3. im Rand von E^t nicht liegen. Indem wir das gleiche Verfahren für jeden Eckpunkt von γ^1 durchführen, übersehen wir kombinatorisch alle 2-dimensionalen im Rande von E^t liegenden Kugeln. Für die höherdimensionalen stellen wir die entsprechende Überlegung an.

Diese kombinatorischen Vorschriften werden bei Anwendung auf Zahlkörper eine wesentliche Rolle spielen; sie sind aber auch an sich von Interesse. In folgenden Arbeiten sollen mit ihrer Hilfe Flächen F^{n-2} in Γ^{n-1} aufgebaut werden, die nicht einer Sphäre, sondern dem $(n-2)$ -dimensionalen Euklidi-

schen Raum homöomorph sind. Auf Grund der oben erwähnten Realisierung von $\Gamma^1, \dots, \Gamma^{n-1}$ in $B(V)$ wird dann unter Vorgabe eines beliebigen Fundamentalbereiches von Γ^1 im Komplex Γ^{n-1} ein $(n-1)$ -dimensionaler Fundamentalbereich einer Deckbewegungsgruppe als Durchschnitt solcher Halbräume in Γ^{n-1} konstruiert, deren Begrenzung in Γ^{n-1} von einer F^{n-2} geliefert werden. Die äquivalenten Seitenpaare des Fundamentalbereichs geben ein Erzeugendensystem der Gruppe, die je nach Wahl von M der gesamten Einheitengruppe des Körpers oder einer Untergruppe isomorph ist.

Im folgenden werden Beispiele für die eben angeführten Sätze 1. bis 3. gebracht.

Dazu einige Vorbemerkungen. Bezüglich einer Menge M im 4-dimensionalen Euklidischen Raum, welche die auf S. 6, 7 gestellten Forderungen erfüllt und außerdem eine dreigliedrige diskrete Gruppe \mathfrak{G} von Streckungen $x' = \varepsilon x$, d. h.

$$x'_1 = \varepsilon_1 x_1 \quad x'_2 = \varepsilon_2 x_2 \quad x'_3 = \varepsilon_3 x_3 \quad x'_4 = \varepsilon_4 x_4$$

gestattet, wurden die extremen Quader (4. Art, vgl. S. 7) bestimmt. Bezeichnen wir zwei extreme Quader, die durch Anwendung eines Elementes von \mathfrak{G} auseinander hervorgehen, als äquivalent, im anderen Falle als inäquivalent, so ergaben sich mod \mathfrak{G} sieben inäquivalente extreme Quader A, B, C, D, E, F, G . Bezeichnen wir einen extremen Quader Q , der aus dem Quader P durch die Streckung $\varepsilon \in \mathfrak{G}$ hervorgeht, mit εP , so läßt sich jeder beliebige extreme Quader (4. Art) eindeutig in den g. r. Exponenten α, β, γ durch $\alpha^a \beta^b \gamma^c H$ darstellen, wo α, β, γ feste Basiselemente von \mathfrak{G} sind und $H \subset A, B, C, D, E, F, G$; und jeder Quader $\alpha^a \beta^b \gamma^c H$, wo α, β, γ, H die eben gegebene Bedeutung haben, ist auch extrem (4. Art). Ist $Q = \varepsilon P$ und Ω_{rs} auf P anwendbar, so gilt $\Omega_{rs}(Q) = \varepsilon \Omega_{rs}(P)$.

In der nun folgenden Tafel, die uns den Graphen Γ^1 und die seinen Strecken zugeordneten Indexpaare liefert, ist der Einfachheit halber für $\Omega_{rs}(P)$ P_{rs} geschrieben.

$$\begin{array}{llll} A_{14} = B, & A_{21} = \gamma^{-1}G, & A_{34} = C, & A_{41} = \gamma^{-1}F, \\ B_{12} = \alpha E, & B_{24} = D, & B_{31} = C, & B_{41} = A, \\ C_{13} = B, & C_{24} = E, & C_{32} = \beta^{-1}D, & C_{43} = A, \\ D_{13} = \alpha G, & D_{23} = \beta C, & D_{34} = F, & D_{42} = B, \\ E_{13} = F, & E_{21} = \alpha^{-1}B, & E_{32} = \beta^{-1}G, & E_{42} = C, \\ F_{14} = \gamma A, & F_{24} = G, & F_{31} = E, & F_{43} = D, \\ G_{12} = \gamma A, & G_{23} = \beta E, & G_{31} = \alpha^{-1}D, & G_{42} = F. \end{array}$$

Nehmen wir einen beliebigen extremen Quader, etwa A , so liegt dieser nach 2. im Rande von $\binom{6}{2} = 4$ 2-dimensionalen „Kugeln“ aus K . Wir wollen diejenige der 6 Kugeln bestimmen, auf deren Rand A , $\Omega_{1k(1)}(A)$,

$\Omega_{2k(2)}(A)$ liegen. Dazu wenden wir 1. an. Wir setzen $l_1 = 1, l_2 = 2$; aus unserer Tafel folgt $k(l_1) = 4, P_1 = B$. Da $k(l_1) = 4 \neq l_1, l_2$ ist, sind die Ziffern l'_1, l'_2 , für welche $\Omega_{l'_1 k(l'_1)}(B), \Omega_{l'_2 k(l'_2)}(B)$ im Rande unserer Kugel liegen, $l'_2 = l_2 = 2, l'_1 = k(l_1) = 4$.

Wir streichen die Akzente und setzen, um die Regel von 1. bequem auf B anwenden zu können, $l_1 = 2, l_2 = 4$. Da $\Omega_{14}(A) = B$, gelangen wir durch Anwendung von $\Omega_{l_2 k(l_2)} = \Omega_{41}$ auf B wieder zurück nach A . Um einen neuen extremen Quader des Randzykels unserer Kugel zu erhalten, setzen wir also $i = 1$. Nach unserer Tafel ist $k(l_1) = 4, \Omega_{24}(B) = D$. Da $k(l_2) \subset (l_1, l_2)$, folgt $l'_1 = 2, l'_2 = 4$.

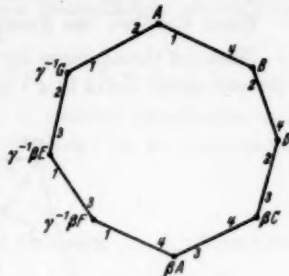


Fig. 1.

Da die weitere Bestimmung des Randzykels immer nach dem gleichen Schema verläuft, wird dabei auf den Text verzichtet.

$$\begin{aligned}
 & D \quad l_1 = 2 \quad l_2 = 4, \\
 & i = 1 \quad k(l_1) = 3 \quad \Omega_{23}(D) = \beta C, \\
 & k(l_2) \subset (l_1, l_2) \quad l'_1 = 3 \quad l'_2 = 4, \\
 & \beta C \quad l_1 = 3 \quad l_2 = 4, \\
 & i = 2 \quad k(l_2) = 3 \quad \Omega_{43}(\beta C) = \beta A, \\
 & k(l_2) \subset (l_1, l_2) \quad l'_1 = 3 \quad l'_2 = 4, \\
 & \beta A \quad l_1 = 3 \quad l_2 = 4, \\
 & i = 2 \quad k(l_2) = 1 \quad \Omega_{41}(\beta A) = \gamma^{-1}\beta F, \\
 & k(l_2) \subset (l_1, l_2) \quad l'_1 = 3 \quad l'_2 = 1, \\
 & \gamma^{-1}\beta F \quad l_1 = 3 \quad l_2 = 1, \\
 & i = 1 \quad k(l_1) = 1 \quad \Omega_{31}(\gamma^{-1}\beta F) = \gamma^{-1}\beta E, \\
 & k(l_1) \subset (l_1, l_2) \quad l'_1 = 3 \quad l'_2 = 1, \\
 & \gamma^{-1}\beta E \quad l_1 = 3 \quad l_2 = 1, \\
 & i = 1 \quad k(l_1) = 2 \quad \Omega_{32}(\gamma^{-1}\beta E) = \gamma^{-1}G, \\
 & k(l_1) \subset (l_1, l_2) \quad l'_1 = 2 \quad l'_2 = 1, \\
 & \gamma^{-1}G \quad l_1 = 2 \quad l_2 = 1, \\
 & i = 2 \quad k(l_2) = 2 \quad \Omega_{12}(\gamma^{-1}G) = A.
 \end{aligned}$$

Damit sind wir zum Ausgangspunkt A zurückgekehrt. Für jeden berechneten Quader Q wurden die Zahlen l_1, l_2 mitbestimmt. Man überzeugt

sich leicht, daß stets $\Omega_{i_1, i_2}(Q), \Omega_{i_2, i_3}(Q)$ in der Menge der ermittelten Quader enthalten sind. Aus der Schlußbemerkung von 1. folgt also, daß wir bereits den ganzen Randzykel bestimmt haben.

Einen Graphen des Randzykels mit den Streckenindizes gibt Fig. 1.

Einen auf Grund des in der Einleitung gegebenen Verfahrens berechneten 2-dimensionalen Zykel in I^3 gibt Fig. 2.

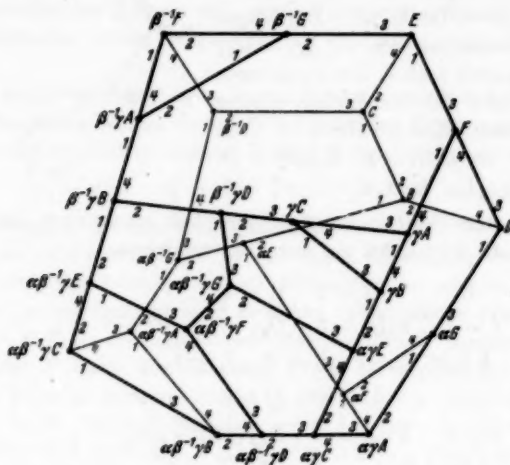


Fig. 2.

Gegeben sei eine Punktmenge M des n -dimensionalen Euklidischen Raumes R_n ($n \geq 3$) mit folgenden Eigenschaften:

1. Alle Punkte von M liegen in dem Teilgebiet O_n

$$x_1 > 0 \dots x_n > 0$$

des R_n .

2. M hat in R_n keine Häufungspunkte.
3. Auf jeder zu einer Koordinatenebene parallelen Ebene liegt höchstens ein Element von M .
4. In jedem Bereich

$$0 < x_{i_1} < c_{i_1} \dots 0 < x_{i_{n-1}} < c_{i_{n-1}}$$

(i_1, \dots, i_n beliebige paarweise verschiedene unter den Zahlen $1, \dots, n$, $c_{i_1}, \dots, c_{i_{n-1}}$ beliebig positiv) liegt mindestens ein Element von M .

Elemente von M bezeichnen wir als M -Punkte.

Wir betrachten Quader

$$0 < x_1 < c_1 \dots 0 < x_n < c_n.$$

Für einen solchen Quader Q definieren wir Seitenflächen verschiedener (0- bis $(n-1)$ -ter) Dimension: Der Bereich

$$x_{s_1} = c_{s_1} \dots x_{s_r} = x_{s_r} \quad 0 < x_{v_1} < c_{v_1} \dots 0 < x_{v_t} < c_{v_t}$$

($r > 0, t \geq 1, r + t = n, s_1, \dots, s_r, v_1, \dots, v_t$ paarweis verschiedene unter den Zahlen $1, \dots, n$) heie (s_1, \dots, s_r) -Seitenflche; sie ist t -dimensional. Der Punkt

$$x_{s_1} = c_{s_1} \dots x_{s_n} = c_{s_n}$$

(s_1, \dots, s_n durchlaufen die Zahlen $1, \dots, n$) heie (s_1, \dots, s_n) -Seitenflche; sie ist 0-dimensional.

Die Seitenflchen von Q liegen smtlich in O_n und machen die Begrenzung von Q in O_n aus. Sie sind punktfremd und gehren Q nicht an.

Wir setzen die Quader in Beziehung zu M . Q heit *frei*, wenn er keinen M -Punkt enthlt.

Ist Q frei, so nehmen wir im Bereich

$$0 < x_m < c_m \quad (m \neq l) \quad 0 < x_l \quad (l \text{ eine der Zahlen } 1, \dots, n)$$

den M -Punkt $\mathfrak{d} = (\delta_1, \dots, \delta_n)$ mit kleinster l -ter Koordinate, der wegen Axiom 2, 3, 4 existiert. Dann verstehen wir unter $d_l(Q)$ den Quader

$$0 < x_m < c_m \quad (m \neq l) \quad 0 < x_l < \delta_l.$$

$d_l(Q)$ ist frei. Er enthlt in der (l) -Seitenflche den M -Punkt \mathfrak{d} . Liegt in der (l) -Seitenflche von Q ein M -Punkt, so ist dieser mit \mathfrak{d} gleich und $d_l(Q) = Q$. Im anderen Falle ist $d_l(Q) \supset Q$.

Sei r eine vorgegebene ganze Zahl im Intervall $1 \leq r \leq n$. Q sei frei und enthalte auf der $(l_{p1}, \dots, l_{pq_p})$ -Seitenflche den M -Punkt m_p ($1 \leq p \leq r, q_1, \dots, q_r \geq 1, q_1 + \dots + q_r = n, l_{v_m}$ paarweis verschiedene unter den Zahlen $1, \dots, n$). Dann heit Q *extrem r -ter Art*. Das Schema

$$(1) \quad (l_{11}, \dots, l_{1q_1}) \dots (l_{r1}, \dots, l_{rq_r})$$

heit *Typus* von Q . Der Typus ist wegen Axiom 3 bis auf Permutation der „Serien“ $(l_{p1}, \dots, l_{pq_p})$ und Permutation der Ziffern innerhalb einer Serie eindeutig durch Q bestimmt.

Ein extremer Quader r -ter Art kann auch dadurch charakterisiert werden, da er frei ist und in jeder in O_n abgeschlossenen (l) -Seitenflche einen M -Punkt enthlt, insgesamt r verschiedene.

Diese Eigenschaften folgen aus der obigen Definition unter Bercksichtigung von Axiom 3. Ist nun umgekehrt ein freier Quader Q vorgelegt, der auf

jeder in O_n abgeschlossenen (l) -Seitenfläche einen M -Punkt enthält, insgesamt r verschiedene, m_1, \dots, m_r , so liege m_p auf den abgeschlossenen $(l_{p_1}), \dots, (l_{p_{q_p}})$ -Seitenflächen, wobei $l_{p_{m_1}} \neq l_{p_{m_2}}$ für $m_1 \neq m_2$ sein soll. Es gilt dann auch $l_{p_1 m_1} \neq l_{p_2 m_2}$ für $p_1 \neq p_2$ und beliebige m_1, m_2 , weil sonst nach Axiom 3 $m_{p_1} = m_{p_2}$ folgte. Da auf jeder (l) -Seitenfläche eines der m_p liegt, kommen unter den $l_{p_{m_i}}$ alle Zahlen $1, \dots, n$ genau einmal vor. Der Durchschnitt der in O_n abgeschlossenen $(l_{p_1}), \dots, (l_{p_{q_p}})$ -Seitenfläche ist der abgeschlossenen $(l_{p_1}, \dots, l_{p_{q_p}})$ -Seitenfläche gleich. Wegen Axiom 3 muß dann m_p in der (offenen) $(l_{p_1}, \dots, l_{p_{q_p}})$ -Seitenfläche liegen.

Es gibt extreme Quader 1., 2., ..., $(n-1)$ -ter Art. Zum Beweise nehmen wir einen M -Punkt $m = (\mu_1, \dots, \mu_n)$, für welchen der Quader Q $0 < x_m < \mu_m$ ($m=1, \dots, n$) frei ist. m existiert wegen Axiom 1, 2. Q ist 1. Art. Ich bilde $d_1(Q)$, $d_2 d_1(Q)$, ..., $d_{n-1} \dots d_1(Q)$ und erhalte sukzessiv Quader 2., 3., ..., n -ter Art.

Ist ein Quader r -ter Art A vorgelegt, so ordnen wir ihm die Matrix (α_{pq}) zu. Hierbei ist $\alpha_m = (\alpha_{m1}, \dots, \alpha_{mn})$ der M -Punkt, der auf der abgeschlossenen (m) -Seitenfläche von A liegt. Wir bezeichnen künftig die M -Punkte eines extremen Quaders, der mit einem lateinischen Buchstaben bezeichnet ist, durch den entsprechenden deutschen Buchstaben, ihre Koordinaten mit den entsprechenden griechischen Buchstaben. In jeder Spalte der Matrix (α_{pq}) ist das Element der Hauptdiagonale maximal. Es gilt ja $\alpha_{pq} \leq c_q$, da a_p auf der Begrenzung von A liegt. Ferner gilt $\alpha_{qq} = c_q$ nach Def. von a_q , woraus die Behauptung folgt. Ist A n -ter Art, so ist α_{qq} das größte Element in der q -ten Spalte. Wir schreiben im folgenden häufig $A = (\alpha_{pq})$ und sprechen vom Quader (α_{pq}) .

Wir definieren jetzt *Intervalle*. Seien $s_1, \dots, s_r, v_1, \dots, v_t$ Zahlen der oben schon erklärten Bedeutung. Ist ein Quader n -ter Art (α_{pq}) vorgelegt, so heißt ein Bereich

$$(2) \quad \alpha_{s_1 s_1} = x_{s_1} \dots \alpha_{s_r s_r} = x_{s_r} \max(\alpha_{s_1 v_1}, \dots, \alpha_{s_r v_t}) \leq x_{v_1} < \alpha_{v_1 v_1} \dots \\ \dots \max(\alpha_{s_1 v_t}, \dots, \alpha_{s_r v_t}) \leq x_{v_t} < \alpha_{v_t v_t}$$

t -Intervall, geschrieben I^t . Ist $t=0$, so fallen die Ungleichungen rechter Hand fort.

Ein I^t ist t -dimensional. Da A n -ter Art ist, gilt $\alpha_{qq} > \alpha_{pq}$ für $p \neq q$, so daß unsere Definition sinnvoll ist. Ferner haben wir

$$\alpha_{s_1 s_1} = \max(\alpha_{s_1 s_1}, \dots, \alpha_{s_r s_1}) \\ \dots \dots \dots \alpha_{s_r s_r} = \max(\alpha_{s_1 s_r}, \dots, \alpha_{s_r s_r}).$$

Wir nennen den Punkt x mit den Koordinaten

$$\xi_m = \max(\alpha_{s_1 m}, \dots, \alpha_{s_r m}) \quad (m=1, \dots, n)$$

Anfangspunkt des Intervalls (2). Er hat die Koordinaten der linken Seiten von (2), ist also der Punkt mit kleinsten Koordinaten im Intervall.

Jede Seitenfläche unseres Quaders n -ter Art enthält eines der oben angeführten Intervalle.

Ein Intervall I^t , das als *Punktmenge* gegeben ist, bestimmt den Quader A n -ter Art, dessen Begrenzung es angehört, und die Seitenfläche von A , in der es liegt, eindeutig. Denn in \bar{I}^t (der in O_n abgeschlossenen Hülle von I^t) liegt als Punkt mit maximalen Koordinaten der Punkt $(\alpha_{11}, \dots, \alpha_{nn})$ der Eckpunkt unseres Quaders A . Die Seitenfläche von A , in welcher I^t liegt, ist dann eindeutig bestimmt durch den Anfangspunkt von I^t , welcher in I^t der Punkt mit kleinsten Koordinaten ist. A heißt dem I^t zugeordnet.

1. Hilfssatz. Der Anfangspunkt eines I^t ist Eckpunkt eines Quaders r -ter Art ($t = n - r$).

Beweis. Für $t = 0$ ist die Behauptung klar. Sei $t \geq 1$. Wir nennen den Quader mit dem Anfangspunkt von I^t als Eckpunkt D . I^t sei der Quader A zugeordnet. Es gilt $D \subset A$, D ist also frei und für die eventuellen M -Punkte auf der Begrenzung von A kommen nur die Punkte a_1, \dots, a_n in Frage. Dabei fallen die Punkte a_{v_1}, \dots, a_{v_t} noch aus wegen

$$\alpha_{v_1, v_1} > \max(\alpha_{s_1, v_1}, \dots, \alpha_{s_r, v_1}).$$

Dagegen liegt der Punkt a_{s_i} ($i = 1, \dots, r$) wegen

$$\begin{aligned} \alpha_{s_i, p} &\leq \max(\alpha_{s_1, p}, \dots, \alpha_{s_r, p}) & p = 1, \dots, n \\ \alpha_{s_i, s_i} &= \max(\alpha_{s_1, s_i}, \dots, \alpha_{s_r, s_i}) \end{aligned}$$

auf der abgeschlossenen (s_i) -Seitenfläche, und auf der abgeschlossenen (v_i) -Seitenfläche ($i = 1, \dots, t$) liegt, wenn etwa $\max(\alpha_{s_1, v_i}, \dots, \alpha_{s_r, v_i}) = \alpha_{s_k, v_i}$ (k eine der Zahlen $1, \dots, r$) ist, wegen $\alpha_{s_k, p} \leq \max(\alpha_{s_1, p}, \dots, \alpha_{s_r, p})$ ($p = 1, \dots, n$) der M -Punkt a_{s_k} . Also liegt auf jeder abgeschlossenen $(n-1)$ -dimensionalen Seitenfläche von D einer der Punkte a_{s_1}, \dots, a_{s_r} , und diese kommen alle vor. D ist also r -ter Art.

Es soll jetzt ein Satz bewiesen werden, der die Umkehrung von Hilfssatz 1 enthält:

2. Hilfssatz. Ist ein Quader D mit der Matrix (δ_{pq}) r -ter Art ($1 \leq r \leq n-1$) vom Typus

$$(l_{11}, \dots, l_{1q_1}), \dots, (l_{r1}, \dots, l_{rq_r})$$

gegeben, ist s_1, \dots, s_r ein beliebiges System von Zahlen, das aus jeder Serie genau ein Element enthält, also etwa $s_p \in (l_{p1}, \dots, l_{pq_p})$, so liegt in dem Bereich

$$(3) \quad x_m = \delta_{mm} \quad (m = s_1, \dots, s_r) \quad x_m \geq \delta_{mm} \quad (m = v_1, \dots, v_t)$$

mindestens ein I^t mit dem Eckpunkt von D als Anfangspunkt.

Sind s_1, \dots, s_r paarweis verschiedene unter den Zahlen $1, \dots, n$, von denen mindestens zwei in derselben Serie des Typus von D liegen, so ist kein I^t mit dem Eckpunkt von D als Anfangspunkt ganz in (3) enthalten.

Beweis. (1. Teil der Beh.). Ich bilde den Quader $C = d_{s_1} \dots d_{s_r}(D)$ mit der Matrix (γ_{sq}) . Er ist n -ter Art. Das I^t

$$x_m = \gamma_{mm} \quad (m = s_1, \dots, s_r) \quad \max_{m=s_1, \dots, s_r} (\gamma_{mq}) \leq x_q < \gamma_{qq} \quad (q = v_1, \dots, v_t)$$

hat den Eckpunkt von D als Anfangspunkt und liegt in (3), denn es gilt

$$c_m = d_m \quad (m = s_1, \dots, s_r)$$

und wegen der besonderen Auswahl der s_m

$$d_q = d_{l(q)} \quad (q = v_1, \dots, v_t, l(q) \subset s_1, \dots, s_r),$$

also

$$\gamma_{mm} = \delta_{mm} \quad (m = s_1, \dots, s_r), \quad \max_{m=s_1, \dots, s_r} (\gamma_{mq}) = \max_{m=s_1, \dots, s_r} (\delta_{mq}) \quad (q = v_1, \dots, v_t),$$

$$\delta_{qq} = \delta_{l(q)q} \leq \max_{m=s_1, \dots, s_r} (\delta_{mq}) \leq \delta_{qq} \quad (q = v_1, \dots, v_t).$$

Wir kommen zum 2. Teil der Behauptung.

Seien s_1, \dots, s_r unter den Zahlen $1, \dots, n$ jetzt so gewählt, daß eine Serie des Typus von D kein s_m enthält. In dieser Serie liegen dann also nur Zahlen v aus dem System der v_1, \dots, v_t .

Sei nun ein Quader n -ter Art $C = (\gamma_{sq})$ angenommen, auf dessen Begrenzung das Intervall

$$x_m = \gamma_{mm} \quad (m = s_1, \dots, s_r) \quad \max_{m=s_1, \dots, s_r} (\gamma_{mq}) \leq x_q < \gamma_{qq} \quad (q = v_1, \dots, v_t)$$

mit dem Eckpunkt von D als Anfangspunkt liegt. Es gilt

$$\gamma_{mm} = \delta_{mm} \quad (m = s_1, \dots, s_r) \quad \max_{m=s_1, \dots, s_r} (\gamma_{mq}) = \delta_{qq} < \gamma_{qq} \quad (q = v_1, \dots, v_t).$$

Wir haben $d_v \neq d_m$ ($m = s_1, \dots, s_r$, v eine der oben definierten Zahlen v) also

$$\delta_{vm} < \delta_{mm} = \gamma_{mm} \quad (m = s_1, \dots, s_r) \quad \delta_{vm} \leq \delta_{mm} < \gamma_{mm} \quad (m = v_1, \dots, v_t).$$

d_v liegt also in C (Widerspruch).

Eine wesentliche Eigenschaft der Intervalle geht hervor aus dem

1. Satz. Die Vereinigungsmenge V aller extremen Quader n -ter Art hat eine Begrenzung $B(V)$ in O_n ; diese wird genau von den Punkten der Intervalle geliefert.

Beweis. Sei etwa das Intervall (2) vorgelegt und $x = (\xi_1, \dots, \xi_n)$ ein beliebiger seiner Punkte. Wäre x Punkt eines extremen Quaders T n -ter Art,

so läge auch der Anfangspunkt von (2) in T , mit ihm der ganze Quader D , der diesen Anfangspunkt als Eckpunkt hat, also auch die r M -Punkte auf der Begrenzung von D . T wäre also nicht frei. x gehört demnach *nicht* zu V . Andererseits liegen in einer beliebigen Umgebung von x Punkte von $A = (\alpha_{pq})$; x ist also Begrenzungspunkt von V .

Die Intervalle bestehen also nur aus Punkten von $B(V)$.

Wir zeigen nunmehr, daß Punkte von $B(V)$ notwendig auf der Begrenzung eines extremen Quaders n -ter Art liegen müssen. Sei etwa x Punkt von $B(V)$. Der Quader $U: 0 < x_m < \xi_m$ ist dann frei. Im anderen Falle enthielte er einen M -Punkt m ; ich wähle dann in der Umgebung

$$|\xi_m - x_m| < \min_m (\xi_m - \mu_m) \quad (m = 1, \dots, n)$$

von x einen Punkt $z = (\zeta_1, \dots, \zeta_n)$ von V . Er gehört einem extremen Quader n -ter Art W an. Alsdann gilt

$$\xi_p - \zeta_p < \xi_p - \mu_p \quad (p = 1, \dots, n),$$

also $\mu_p < \zeta_p$, $m \in W$ (Widerspruch).

Ich bilde jetzt $d_1 d_2 \dots d_n(U)$, der ein Quader n -ter Art ist und auf seiner Begrenzung x enthält.

Der Beweis von Satz 1 wird geliefert sein, wenn wir zeigen: Ein Punkt p der Begrenzung eines Quaders $A = (\alpha_{pq})$ n -ter Art, der keinem Intervall angehört, dem A zugeordnet ist, gehört notwendig zu V .

Die Dimension t der Seitenfläche von A , welche p enthält, muß ≥ 1 sein, da der Eckpunkt von A ein I^0 ist, dem p nicht angehören soll. Sei also die p enthaltende Seitenfläche von A durch

$$x_m = \alpha_{m m} \quad (m = s_1, \dots, s_r) \quad 0 < x_m < \alpha_{m m} \quad (m = v_1, \dots, v_t)$$

gegeben. Auf ihr liegt das t -Intervall

$$x_m = \alpha_{m m} \quad (m = s_1, \dots, s_r) \quad \max_{m = s_1, \dots, s_r} (\alpha_{m q}) \leq x_q < \alpha_{q q} \quad (q = v_1, \dots, v_t).$$

Der Rest unserer t -dimensionalen Seitenfläche ist im Falle $t \geq 2$ die Vereinigungsmenge der t Stücke

$$x_m = \alpha_{m m} \quad (m = s_1, \dots, s_r) \quad 0 < x_q < \max_{m = s_1, \dots, s_r} (\alpha_{m q})$$

$$0 < x_m < \alpha_{m m} \quad (m \neq s_1, \dots, s_r, q),$$

wobei jedes $q \in (v_1, \dots, v_t)$ ein Stück liefert, im Falle $t = 1$ das Stück

$$x_m = \alpha_{m m} \quad (m = s_1, \dots, s_r) \quad 0 < x_{v_1} < \max_{m = s_1, \dots, s_r} (\alpha_{m v_1}).$$

Wir können ohne Beschränkung der Allgemeinheit annehmen, daß p in dem durch $q = v_1$ gegebenen Stück liegt (Permutation der v_i). Wir untersuchen den Quader Q

$$0 < x_m < \alpha_{m v_1} \quad (m \neq v_1) \quad 0 < x_{v_1} < \max_{m=s_1, \dots, s_r} (\alpha_{m v_1}),$$

der als Teilquader von A frei ist.

Sei $r \geq 2$. Wir dürfen annehmen, daß

$$(4) \quad \max_{m=s_1, \dots, s_r} (\alpha_{m v_1}) = \alpha_{s_1 v_1}$$

ist (Permutation der s_i). Q enthält auf seiner (s_2) -, \dots , (s_r) -Seitenfläche bzw. die Punkte a_{s_2}, \dots, a_{s_r} . Die (s_1) -Seitenfläche ist frei. Die $(s_1 v_1)$ -Seitenfläche enthält a_{s_1} . Wir nehmen den Quader $B = d_{s_1} d_{v_2} \dots d_{v_t}(Q)$, im Falle $t = 1$ $B = d_{s_1}(Q)$. Dieser Quader enthält auf der (s_1) -, (v_2) -, \dots , (v_t) -Seitenfläche bzw. die M -Punkte $b_{s_1}, b_{v_2}, \dots, b_{v_t}$. Es gilt

$$\beta_{s_1 s_1} > \alpha_{s_1 s_1} \quad \beta_{v_2 v_2} \geq \alpha_{v_2 v_2} \dots \beta_{v_t v_t} \geq \alpha_{v_t v_t}.$$

Außerdem gilt $a_{s_1} = b_{v_1}$, $a_{s_2} = b_{s_2}$, \dots , $a_{s_r} = b_{s_r}$. Da die Seitenflächen punktfremd sind, enthält der freie B auf jeder abgeschlossenen $(n-1)$ -dimensionalen Seitenfläche einen M -Punkt, insgesamt n Stück, ist also n -ter Art. Für unseren Punkt $p = (\pi_1, \dots, \pi_n)$ gilt dann

$$0 < \pi_{s_1} = \alpha_{s_1 v_1} < \beta_{s_1 s_1} \quad \pi_{s_2} = \beta_{s_2 s_2} \dots \pi_{s_r} = \beta_{s_r s_r}$$

$$0 < \pi_{v_1} < \max_{m=s_1, \dots, s_r} (\alpha_{m v_1}) = \alpha_{s_1 v_1} = \beta_{s_1 v_1}$$

$$0 < \pi_{v_2} < \alpha_{v_2 v_2} \leq \beta_{v_2 v_2} \dots 0 < \pi_{v_t} < \alpha_{v_t v_t} \leq \beta_{v_t v_t},$$

d. h. p liegt in der (s_2, \dots, s_r) -Seitenfläche von B .

Sei $r = 1$. Dann liegt p in $B = d_{s_1} d_{v_2} \dots d_{v_t}(Q)$, ist also Punkt von V . Im Falle $t = n-1$ sind wir also fertig.

Im Falle $t < n-1$ gibt es, wie wir sahen, eine $(t+1)$ -dimensionale Seitenfläche eines extremen Quaders B , in welcher p enthalten ist. Wenden wir nun die gleiche Überlegung auf diese Seitenfläche von B an, wie eben auf die von A , usf., so erhalten wir durch Induktion die Behauptung.

In Satz 2 wird eine weitere wichtige Eigenschaft der Intervalle bewiesen. Wir benötigen dazu den

3. Hilfssatz. Die (s_1, \dots, s_r) -Seitenfläche eines Quaders n -ter Art enthält genau einen Eckpunkt eines extremen Quaders. Dieser ist r -ter Art.

Beweis. Sei der vorgegebene Quader n -ter Art $A = (\alpha_{pq})$. Daß in der (s_1, \dots, s_r) -Seitenfläche von A wenigstens ein Eckpunkt eines Quaders r -ter Art liegt, folgt daraus, daß sie ein t -dimensionales Intervall enthält, also auch dessen Anfangspunkt, der nach Hilfssatz 1 Eckpunkt eines Quaders r -ter Art ist.

Angenommen, auf der (s_1, \dots, s_r) -Seitenfläche von A ($r \leq n-1$)

$$x_m = \alpha_{mm} \quad (m = s_1, \dots, s_r) \quad 0 < x_m < \alpha_{mm} \quad (m = v_1, \dots, v_t)$$

liegt der Eckpunkt des Quaders (π_{pq}) . Dann gilt $p_m = \alpha_m$ ($m = s_1, \dots, s_r$), der Quader (π_{pq}) wird also gegeben durch

$$0 < x_m < \alpha_{mm} \quad (m = s_1, \dots, s_r) \quad 0 < x_m < \pi_{mm} \quad (m = v_1, \dots, v_t).$$

Die Punkte a_{v_1}, \dots, a_{v_t} sind wegen $\pi_{v_m v_m} < \alpha_{v_m v_m}$ nicht auf der Begrenzung von (π_{pq}) gelegen. Da $(\pi_{pq}) \subset A$, kommen also für die Begrenzung von (π_{pq}) nur die M -Punkte a_{s_1}, \dots, a_{s_r} in Frage.

Also kann (π_{pq}) nicht $r' > r$ -ter Art sein. Ist (π_{pq}) $r' \leq r$ -ter Art, so können wir annehmen, daß die M -Punkte $a_{s_1}, \dots, a_{s_{r'}}$ auf seiner Begrenzung liegen. Für die Koordinaten des Eckpunktes von (π_{pq}) folgt $\pi_{mm} = \max(\alpha_{s_1 m}, \dots, \alpha_{s_{r'} m})$ ($m = 1, \dots, n$). Im Falle $r' = r$ ist also dieser Eckpunkt eindeutig als Anfangspunkt des schon genannten I^r bestimmt. Im Falle $r' < r$ liegt er gar nicht in der (s_1, \dots, s_r) -Seitenfläche von A wegen

$$\alpha_{s_1 s_{r'}}, \dots, \alpha_{s_{r'} s_{r'}} < \alpha_{s_r s_r}, \quad \text{also} \quad \pi_{s_r s_r} < \alpha_{s_r s_r},$$

so daß dieser Fall ausscheidet.

2. Satz. *Intervalle verschiedener Dimension sind punktfremd. Intervalle gleicher Dimension haben (dann und) nur dann einen Punkt gemein, wenn sie den gleichen Anfangspunkt haben.*

Beweis. Seien I^{t_1}, K^{t_2} zwei Intervalle. Sei $r_1 = n - t_1$, $r_2 = n - t_2$, $r_1 \geq r_2$. K^{t_2} liege auf der (s_1, \dots, s_{r_2}) -Seitenfläche des zugeordneten Quaders Q_2 . Liegt nun ein Punkt x in I^{t_1} und K^{t_2} , so liegt der Anfangspunkt b von I^{t_1} in einer $(s'_1, \dots, s'_{r'_1})$ -Seitenfläche von Q_2 , wobei $(s'_1, \dots, s'_{r'_1}) \subset (s_1, \dots, s_{r_2})$ und $r'_2 \leq r_2$, oder in Q_2 . Nun ist aber b Eckpunkt eines Quaders r_1 -ter Art nach Hilfssatz 1, kann also nicht in Q_2 liegen. Ist nun $t_1 \neq t_2$, also $r'_2 < r_1$, so kann b wegen Hilfssatz 3 auch nicht in der $(s'_1, \dots, s'_{r'_1})$ -Seitenfläche von Q_2 liegen; der Fall $t_1 \neq t_2$ ist also ausgeschlossen. Ist $t_1 = t_2$, so folgt $r_1 = r_2$. Wegen Hilfssatz 3 muß also das System $(s'_1, \dots, s'_{r'_1})$ (bis auf Permutation der Elemente) mit (s_1, \dots, s_{r_2}) übereinstimmen, und b muß mit dem Anfangspunkt von K^{t_2} identisch sein.

Fassen wir die Intervalle mit gleichem Anfangspunkt zusammen, so erhalten wir nach Satz 1, 2 eine Zerlegung von $B(V)$ in punktfremde Mengen der Dimensionen 0 bis $(n-1)$; wir definieren dementsprechend: Jede Vereinigungsmenge t -dimensionaler Intervalle mit gleichem Anfangspunkt heißt t -Ecke, geschrieben E^t . Den gemeinsamen Anfangspunkt nennen wir Spitze von E^t .

Die t -Ecken sind topologische t -dimensionale offene Kugeln. Der Beweis hierfür sowie für die in der Einleitung erwähnten Eigenschaften dieser Kugeln soll jetzt vorbereitet werden.

4. Hilfssatz. *Die Quader, welche den Intervallen einer festen t -Ecke zugeordnet sind, liegen in einem beschränkten Raumteil. (Folglich enthält jede t -Ecke nur endlich viele t -Intervalle.)*

Beweis. Für $t = 0$ ist der Satz klar, da es sich hier nur um einen Quader handelt. Sei $t \geq 1$. Die Spitze unserer t -Ecke ist Eckpunkt eines Quaders Q r -ter Art. Q ist in jedem Quader S enthalten, der einem t -Intervall unserer t -Ecke zugeordnet ist. Q enthält wegen $r \geq 1$ mindestens einen M -Punkt a auf seiner Begrenzung. Der Quader $A = (x_{pa})$ mit a als Eckpunkt ist in Q enthalten und 1. Art. Der Quader A liegt also in jedem $S = (\sigma_{pa})$. Folglich gilt $\alpha_{mm} \leq \sigma_{mm}$ ($m = 1, \dots, n$).

Der Quader $d_l(A)$ ($l = 1, \dots, n$) enthalte in seiner (l)-Seitenfläche den M -Punkt $b_l = (\delta_{l1}, \dots, \delta_{ln})$. Dann gilt $0 < \delta_{lm} < \alpha_{mm}$ ($m \neq l$). Es gibt nun kein S mit $\sigma_{ll} > \delta_{ll}$ (l eine beliebige der Zahlen $1, \dots, n$). Sonst folgte, daß b_l in S gelegen wäre.

Alle S liegen daher im Bereich $0 < x_m \leq \delta_{mm}$ ($m = 1, \dots, n$), in einem beschränkten Raumteil. Nach Axiom 2 gibt es also nur endlich viele S , daher in E^t nur endlich viele I^t .

5. Hilfssatz. *Seien k Serien*

$$(l_{11}, \dots, l_{1q_1}) \dots, (l_{k1}, \dots, l_{kq_k})$$

von Zahlen l_{pq} gegeben, die paarweis verschieden unter den Zahlen $1, \dots, n$ gewählt sind, $q_1 \geq 2$, $\lambda = q_1 + \dots + q_k - k$. Sei e_m der Einheitspunkt auf der m -ten Achse des Koordinatensystems ($m = 1, \dots, n$).

Der Komplex, der aus allen $(\lambda - 1)$ -Simplexen $S_{v_1, \dots, v_{\lambda-1}}$ mit den Eckpunkten $e_{v_1}, \dots, e_{v_{\lambda-1}}$ besteht, wo $v_1, \dots, v_{\lambda-1}$ ein System von Zahlen bedeutet, das aus der m -ten der obigen Serien ($m = 1, \dots, k$) $q_m - 1$ Elemente enthält, ist homöomorph der $(\lambda - 1)$ -dimensionalen Sphäre.

Beispiel. Sei $k = 2$, $q_1 = 2$, $q_2 = 1$, $l_{11} = 3$, $l_{12} = 1$, $l_{21} = 2$. Es ist $\lambda = 1$. Wir haben die Serien (31) (2). Sie definieren den Komplex, der aus dem Punktepaar e_3, e_1 besteht, eine 0-dimensionale Sphäre.

Sei $k = 1$, $q_1 = 3$, $l_{11} = 1$, $l_{12} = 2$, $l_{13} = 3$. Es ist $\lambda = 2$. Wir haben die Serie (123). Der zugeordnete Komplex besteht aus den Strecken e_1e_2 , e_1e_3 , e_2e_3 , die einen Dreiecksrand, eine 1-Sphäre ergeben.

Der Beweis des Satzes erfolgt durch Induktion nach k .

Jedes Simplex S_{v_1, \dots, v_1} ist in dem Komplex enthalten, der aus dem von den e_1, \dots, e_n aufgespannten $(n-1)$ -Simplex und dessen Teilsimplexen besteht. Also ist $\Sigma S_{v_1, \dots, v_1}$ ein Komplex.

Für $k=1$ und ein beliebiges $q_1 \geq 2$ ist die Behauptung richtig. In diesem Fall besteht unser Komplex aus dem Rande des (q_1-1) -Simplexes mit den Ecken $e_{1,1}, \dots, e_{1,q_1}$.

Angenommen, die Behauptung wäre für ein $k=k' \geq 1$ richtig. Dann ist also der Komplex $\Sigma S_{v_1, \dots, v_2} (\lambda' = q_1 + \dots + q_{k'} - k')$, welcher den Serien $(l_{1,1}, \dots, l_{1,q_1}) \dots (l_{k',1}, \dots, l_{k',q_{k'}})$ zugeordnet ist, eine $(\lambda' - 1)$ -Sphäre.

Wir wollen die Behauptung für $k=k'+1=k''$ beweisen. Sei $\lambda'' = q_1 + \dots + q_{k'} + q_{k''} - k''$. Wir nehmen die Kette der Simplexe mit den Eckpunkten

$$(5) \quad e_{v_1}, \dots, e_{v_2}, e_{1_{k''},1}, \dots, e_{1_{k''},q_{k''}}.$$

Sie entsteht durch Projektion der Sphäre $\Sigma S_{v_1, \dots, v_2}$ vom Punkte $e_{1_{k''},1}$ aus, wodurch wir eine topologische λ' -dimensionale Kugel erhalten, durch weitere Projektion dieser Kugel von $e_{1_{k''},2}$ aus, wodurch eine $(\lambda' + 1)$ -dimensionale Kugel entsteht usf. Es folgt, daß die Kette (5) der λ'' -dimensionalen Kugel homöomorph ist. Wir bilden jetzt den Rand der Kette (5) mod 2. Er ist die Summe der Ränder der Simplexe (5) mod 2. Diese Summe enthält die $(\lambda'' - 1)$ -dimensionalen Simplexe des Komplexes, welcher

$$(l_{1,1}, \dots, l_{1,q_1}) \dots (l_{k'',1}, \dots, l_{k'',q_{k''}})$$

zugeordnet ist, jedes genau einmal, die übrigen $(\lambda'' - 1)$ -dimensionalen Simplexe in gerader Anzahl, da der Rand der Sphäre $\Sigma S_{v_1, \dots, v_2}$ mod 2 verschwindet. Für $k=k'+1$ ist also unser Komplex der Rand einer λ'' -dimensionalen Kugel, also $(\lambda'' - 1)$ -Sphäre, w. z. b. w.

Die Spitze einer t -Ecke ist nach Hilfssatz 1 Eckpunkt eines Quaders r -ter Art $(\delta_{p,q})$. Die t -Intervalle von E^t liegen nach Hilfssatz 2 in den Ebenen

$$x_{s_1} = \delta_{s_1 s_1} \dots x_{s_r} = \delta_{s_r s_r};$$

hierbei sind die Systeme (s_1, \dots, s_r) bis auf Permutation ihrer Elemente genau die mit $s_p \subset (l_{p,1}, \dots, l_{p,q_p})$, wenn $(l_{1,1}, \dots, l_{1,q_1}) \dots (l_{r,1}, \dots, l_{r,q_r})$ der Typus des Quaders $(\delta_{p,q})$ ist.

6. Hilfssatz. Sei $r \leq n-1$. Der Durchschnitt eines Bereiches

$$H_{s_1, \dots, s_r}: x_{s_1} = \delta_{s_1 s_1} \dots x_{s_r} = \delta_{s_r s_r} \quad x_{v_1} \geq \delta_{v_1 v_1} \dots x_{v_t} \geq \delta_{v_t v_t}$$

mit E^t ist die Vereinigungsmenge aller in H_{s_1, \dots, s_r} gelegenen t -Intervalle von E^t .

Beweis. $\mathbf{x} = (\xi_1, \dots, \xi_n)$ gehöre zu H_{s_1, \dots, s_r} und E^t . Als Punkt von E^t gehört er einem t -Intervall I^t an, und es gilt

$$\xi_{s'_m} = \beta_{s'_m s'_m} \quad (m = 1, \dots, r) \quad \max_{m=1, \dots, r} (\beta_{s'_m s'_m}) \leq \xi_{s'_q} < \beta_{s'_q s'_q} \\ (q = 1, \dots, t).$$

Wir wollen annehmen, daß die Mengen der s_m, s'_m nicht übereinstimmen, andererseits wenigstens ein Element gemein haben, und die gleiche Voraussetzung an die Mengen der v_m, v'_m stellen. Der Beweis verläuft in den hierbei nicht berücksichtigten Fällen ganz analog.

s_m, s'_m, v_m, v'_m seien so geordnet, daß

$$\begin{aligned} s'_1 &= v_1 \dots s'_p = v_p & s'_{p+1} &= s_{p+1} \dots s'_r = s_r, \\ s_1 &= v'_1 \dots s_p = v'_p & v'_{p+1} &= v_{p+1} \dots v'_t = v_t \end{aligned}$$

gilt. Da der Anfangspunkt von I^t mit dem Eckpunkt von $(\delta_{p,q})$ übereinstimmt, folgt

$$(6) \quad \beta_{s'_m s'_m} = \delta_{s_m s_m} \quad (m = p+1, \dots, r),$$

$$(7) \quad \beta_{s'_m s'_m} = \delta_{v_m v_m} \quad (m = 1, \dots, p).$$

Wir nehmen den Quader R

$$\begin{aligned} 0 < x_{s_m} < \delta_{s_m s_m} \quad (m = 1, \dots, r) & \quad 0 < x_{v_m} < \delta_{v_m v_m} \quad (m = 1, \dots, p) \\ 0 < x_{s_m} < \beta_{s_m s_m} \quad (m = p+1, \dots, t). \end{aligned}$$

Er ist Teilquader von $(\beta_{p,q})$ wegen (6), (7) und

$$\delta_{s_m s_m} = \xi_{s_m} = \xi_{s'_m} < \beta_{s'_m s'_m} = \beta_{s_m s_m} \quad (m = 1, \dots, p).$$

R ist also frei. Er enthält die M -Punkte

$$d_{s_m} \quad (m = 1, \dots, r) \quad \text{auf der abgeschl. } (s_m)\text{-Seitenfläche,}$$

$$d_{v_m} \quad (m = 1, \dots, p) \quad \text{auf der abgeschl. } (v_m)\text{-Seitenfläche.}$$

Wegen $d_{v_m} = d_{v_{t(m)}}$ ($m = 1, \dots, n$) (vgl. die Voraussetzungen über die Auswahl der s_1, \dots, s_r) liegt also der M -Punkt d_{v_m} ($m = 1, \dots, p$) in der abgeschl. $(v_m, s_{t(m)})$ -Seitenfläche von R . Infolgedessen gilt

$$(8) \quad R \subseteq d_{v_t}(R) \subseteq \dots \subseteq d_{v_{p+1}} \dots d_{v_t}(R) \subset d_{v_p} \dots d_{v_t}(R) \subset \dots \subset d_{v_1} \dots d_{v_t}(R) = T.$$

$T = (\tau_{p,q})$ enthält auf der abgeschl. (s_m) -Seitenfläche ($m = 1, \dots, r$) den M -Punkt d_{s_m} , in der (v_m) -Seitenfläche die notwendig von den d_{s_m} und untereinander verschiedenen M -Punkte t_{v_m} ($m = 1, \dots, t$). T ist frei, also extrem n -ter Art. Ich behaupte, \mathbf{x} liegt in dem Intervall

$$(9) \quad x_{s_m} = \tau_{s_m s_m} \quad (m = 1, \dots, r) \quad \max_{m=1, \dots, r} (\tau_{s_m s_m}) \leq x_{s'_q} < \tau_{s'_q s'_q} \quad (q = 1, \dots, t).$$

Wie schon gesagt, ist $t_{s,m} = b_{s,m}$ ($m = 1, \dots, r$), d. h. (9) hat den Eckpunkt von $(\delta_{p,q})$ als Anfangspunkt, gehört also zu E^t ; ferner gilt mit Rücksicht auf (7), (8)

$$\begin{aligned}\xi_{v,m} &= \delta_{v,m} v_m < \tau_{v,m} v_m & (m = 1, \dots, p), \\ \xi_{v,m} &< \beta_{v,m} v_m \leq \tau_{v,m} v_m & (m = p+1, \dots, t).\end{aligned}$$

x gehört also dem in H_{s_1, \dots, s_r} gelegenen Intervall (9) von E^t an, womit der Satz bewiesen ist.

Wir nehmen jetzt das Stück unserer t -Ecke, das ganz in H_{s_1, \dots, s_r} liegt, wo (s_1, \dots, s_r) wie eben ein System von Zahlen darstellt, das nach Hilfssatz 2 aus dem Typus des Quaders $(\delta_{p,q})$ entnommen ist. Sei $b = (\delta_1, \dots, \delta_n)$ der Eckpunkt von $(\delta_{p,q})$.

Sei auch im folgenden $r \leq n-1$.

Jeder in H_{s_1, \dots, s_r} liegende Strahl mit b als Anfangspunkt wird gegeben durch

$$(10) \quad b + l(b - b) \quad 0 \leq l < \infty,$$

wo $b = (\beta_1, \dots, \beta_n)$ ein Punkt in H_{s_1, \dots, s_r} mit $\sum_m (\beta_m - \delta_m) = 1$ ist. Nehmen wir irgendein in H_{s_1, \dots, s_r} gelegenes I' , so gibt es ein positives l_0 , für welches $b + l_0(b - b)$ in I' liegt, d. h.

(11) auf jedem Strahl (10) in H_{s_1, \dots, s_r} gibt es einen Punkt von E^t mit $l = l_0 > 0$.

Ferner gilt

(12) Gehört ein Punkt eines festen Strahles (10) mit $l = l_1$ zu E^t , so auch alle Punkte des Strahles mit $0 \leq l < l_1$.

Dieser Punkt, p , gehört als Punkt von E^t einem t -Intervall mit dem Anfangspunkt b an. Die Verbindungsstrecke des Punktes p mit b liegt im gleichen Intervall, also sind alle Punkte von (10) mit $0 \leq l < l_1$ Punkte von E^t .

(13) Für jeden Strahl (10) gibt es ein l_2 , so daß alle Punkte des Strahles mit $l_2 > l$ E^t nicht angehören.

Dies folgt aus Hilfssatz 4.

Aus (11), (12), (13) folgt, daß für jeden in H_{s_1, \dots, s_r} gelegenen Strahl (10) die Werte l , welche Punkte von E^t liefern, eine obere Grenze $l^*(b) > 0$ haben. Alle Punkte mit $0 \leq l < l^*$ gehören zu E^t , die Punkte mit $l > l^*$ gehören nicht zu E^t .

7. Hilfssatz. Die in H_{s_1, \dots, s_r} gelegenen oberen Grenzen $b^* = b + l^*(b)(b - b)$ bilden eine im R_n abgeschlossene Punktmenge.

Beweis. Sei ein Punkt $p = (\pi_1, \dots, \pi_n)$ in H_{s_1, \dots, s_r} gegeben, in dessen Nachbarschaft unendlich viele obere Grenzen $b^* \subset H_{s_1, \dots, s_r}$ liegen. Notwendig ist $p \neq b$. Es folgt

1. Auf der Strecke bp gibt es nur Punkte von E^t .

Sei ein Punkt $r = (\varrho_1, \dots, \varrho_n) \neq b$ dieser Strecke gegeben. Seien v_1, \dots, v_p die Indizes, für welche $\pi_{v_m} = \varrho_{v_m} = \delta_{v_m}$ gilt (den Fall, wo kein v_m existiert, behandelt man mit der gleichen Methode). Sei $\varepsilon < \pi_{v_m} - \varrho_{v_m}$ ($m = p+1, \dots, n$). Eine ε -Nachbarschaft von p in H_{s_1, \dots, s_r} wird dann gegeben durch

$$\begin{aligned} x_{s_1} = \delta_{s_1} \dots x_{s_r} = \delta_{s_r} \quad \delta_{v_1} \leq x_{v_1} < \delta_{v_1} + \varepsilon \dots \delta_{v_p} \leq x_{v_p} < \delta_{v_p} + \varepsilon \\ \pi_{v_m} - \varepsilon \leq x_{v_m} < \pi_{v_m} + \varepsilon \quad (m = p+1, \dots, n). \end{aligned}$$

In dieser Umgebung gibt es eine $b^* = (\beta_1^*, \dots, \beta_n^*)$, welche wegen Hilfsatz 4, 6 der abgeschlossenen Hülle eines in H_{s_1, \dots, s_r} gelegenen I^t angehört. Der I^t zugeordnete Quader sei (γ_{v_i}) . Es gilt:

$$\beta_{s_1}^* = \delta_{s_1} \dots \beta_{s_r}^* = \delta_{s_r} \quad \delta_{v_1} \leq \beta_{v_1}^* < \gamma_{v_1, v_1} \dots \delta_{v_t} \leq \beta_{v_t}^* < \gamma_{v_t, v_t}.$$

Nun ist $\beta_{v_m}^* > \pi_{v_m} - \varepsilon > \varrho_{v_m}$ ($m = p+1, \dots, n$), d. h. r gehört dem I^t an, also E^t , w. z. b. w.

2. Auf der Verlängerung der Strecke bp über p hinaus gibt es keinen Punkt r von E^t .

r müßte sonst einem I^t von E^t angehören, von dem wir nach Hilfsatz 6 annehmen können, daß es in H_{s_1, \dots, s_r} liegt. Auch p muß diesem Intervall angehören. Es gilt also

$$\pi_{s_1} = \delta_{s_1} \dots \pi_{s_r} = \delta_{s_r} \quad \delta_{v_1} \leq \pi_{v_1} < \gamma_{v_1, v_1} \dots \delta_{v_t} \leq \pi_{v_t} < \gamma_{v_t, v_t}.$$

Hierbei ist (γ_{v_i}) der I^t zugeordnete Quader. Seien v_1, \dots, v_t die Indizes, für welche $\pi_{v_m} = \varrho_{v_m} = \delta_{v_m}$ gilt. Sei $\varepsilon < \gamma_{v_m, v_m} - \pi_{v_m}$ ($m = 1, \dots, t$) und $< \pi_{v_m} - \delta_{v_m}$ ($m = p+1, \dots, t$). Eine ε -Nachbarschaft von p in H_{s_1, \dots, s_r} wird dann gegeben durch

$$\begin{aligned} x_{s_1} = \delta_{s_1} \dots x_{s_r} = \delta_{s_r} \quad \delta_{v_1} \leq x_{v_1} < \delta_{v_1} + \varepsilon \dots \delta_{v_p} \leq x_{v_p} < \delta_{v_p} + \varepsilon \\ \pi_{v_m} - \varepsilon < x_{v_m} < \pi_{v_m} + \varepsilon \quad (m = p+1, \dots, t). \end{aligned}$$

Für eine in einer solchen ε -Nachbarschaft liegende obere Grenze b^* gilt dann

$$\beta_{v_m}^* < \delta_{v_m} + \varepsilon = \pi_{v_m} + \varepsilon < \gamma_{v_m, v_m} \quad (m = 1, \dots, p),$$

$$\beta_{v_m}^* < \pi_{v_m} + \varepsilon < \gamma_{v_m, v_m} \quad (m = p+1, \dots, t).$$

b^* liegt also in I^t , ist also nicht obere Grenze, so daß 2. bewiesen ist. 1. und 2. ergeben, daß p der Menge der oberen Grenzen angehört, diese Menge also abgeschlossen ist.

Wir bilden jetzt die Menge G_{s_1, \dots, s_r} der Punkte $b \subset H_{s_1, \dots, s_r}$ mit $\sum_m (\beta_m - \delta_m) = 1$ auf die Menge G_{s_1, \dots, s_r}^* der in H_{s_1, \dots, s_r} gelegenen oberen Grenzen ab, indem wir b die auf dem Strahl von b durch b gelegene obere

Grenze b^* zuordnen. Diese Abbildung A_{s_1, \dots, s_r} ist eineindeutig. Da die b^* in der abgeschlossenen Hülle von E^t liegen, folgt aus Hilfssatz 4, daß G_{s_1, \dots, s_r}^* beschränkt ist. Aus Hilfssatz 7 folgt die Abgeschlossenheit von G_{s_1, \dots, s_r}^* . Die gleichen Eigenschaften hat $\bar{G}_{s_1, \dots, s_r}$. Insgesamt folgt, daß A_{s_1, \dots, s_r} topologisch ist.

Wir definieren nun in der Vereinigungsmenge aller $G_{s_1, \dots, s_r}, \Sigma G_{s_1, \dots, s_r}$, — (s_1, \dots, s_r) durchläuft alle in den Voraussetzungen zu Hilfssatz 6 zugelassenen Systeme — eine eineindeutige Abbildung A durch die Forderung, daß $A = A_{s_1, \dots, s_r}$ in G_{s_1, \dots, s_r} sein soll. Da die G_{s_1, \dots, s_r} abgeschlossene Teilmengen von $\Sigma G_{s_1, \dots, s_r}$ sind und A in jeder dieser Teilmengen topologisch ist, so ist A auch in $\Sigma G_{s_1, \dots, s_r}$ topologisch. Nun ist $\Sigma G_{s_1, \dots, s_r}$ für $r \leq n-1$, d. h. $t \geq 1$, gerade ein Komplex der in Hilfssatz 5 betrachteten Art ($q_1 \geq 2$ läßt sich durch Permutation der Serien des Typus von $(\delta_{p,q})$ erreichen; $k=r, \lambda=n-r=t$), also der $(t-1)$ -Sphäre homöomorph. Daraus folgt: Die Menge der in allen H_{s_1, \dots, s_r} definierten oberen Grenzen b^* ist homöomorph der $(t-1)$ -Sphäre.

Man zeigt nun leicht, daß kein b^* zu E^t gehört. Es müßte sonst in einem I^t von E^t liegen; dann wären noch Punkte der Verlängerung der Strecke bb^* über b^* hinaus zu E^t gehörig, was nicht sein kann. Verbinden wir jetzt b mit jedem Punkt von $\Sigma G_{s_1, \dots, s_r}$, so ist der dadurch entstehende Kegel eine t -dimensionale Vollkugel; bilden wir jede Strecke bb auf bb^* , wo $b^* = A(b)$, linear ab, so daß dem Punkt b der Punkt b entspricht, und beachten wir, daß jeder Strahl von b aus durch ein b genau einen Punkt von $\Sigma G_{s_1, \dots, s_r}^*$ enthält, so erhalten wir eine topologische Abbildung der t -dimensionalen Vollkugel auf $E^t + \Sigma G_{s_1, \dots, s_r}^*$. Da $\Sigma G_{s_1, \dots, s_r}^*$ hierbei auf den Rand der Vollkugel abgebildet wird und keinen Punkt von E^t enthält, gilt der

3. Satz. Jede t -Ecke ($t \geq 1$) ist homöomorph der t -dimensionalen offenen Kugel.

Der Rand von E^t wird von der Menge der oberen Grenzen b^* gebildet, eine Aussage, die wir im 4. Satz in eine andere Form bringen werden. Dazu noch eine Definition: So, wie wir jedem Intervall einen Quader n -ter Art zugeordnet haben, ordnen wir jetzt jedem I^t Intervalle I' höherer Dimension zu: Sei ein I^t ($0 \leq t \leq n-2$) auf dem Quader n -ter Art $(\beta_{p,q})$ vorgelegt:

$$\max(\beta_{s_1, m}, \dots, \beta_{s_r, m}) = x_m \quad (m = s_1, \dots, s_r),$$

$$\max(\beta_{s_1, m}, \dots, \beta_{s_r, m}) \leq x_m < \beta_{m, m} \quad (m = v_1, \dots, v_t).$$

Sei $(s'_1, \dots, s'_{r'})$ in (s_1, \dots, s_r) enthalten, $1 \leq r' < r$, $r' + t' = n$; $v'_1, \dots, v'_{t'}$ seien unter den Zahlen $1, \dots, n$ die von $s'_1, \dots, s'_{r'}$ verschiedenen. Dann heißt das Intervall $I'^{t'}$

$$\max(\beta_{s'_1, m}, \dots, \beta_{s'_{r'}, m}) = x_m \quad (m = s_1, \dots, s_{r'}),$$

$$\max(\beta_{s'_1, m}, \dots, \beta_{s'_{r'}, m}) \leq x_m < \beta_{m, m} \quad (m = v'_1, \dots, v'_{t'})$$

dem Intervall I^t zugeordnet.

Alle I^t zugeordneten I^t liegen auf demselben Quader $(\beta_{n,t})$; I^t liegt in der abgeschlossenen Hülle jedes I^t . Jedem I^t sind $\binom{n-t}{n-t}$ t' -dimensionale Intervalle zugeordnet. Sie sind punktfremd.

Der folgende Satz 4 steht in Parallele zu Satz 1. Die Intervalle unserer t -Ecke entsprechen den Quadern n -ter Art, die t -Ecke der Vereinigungsmenge V aller Quader n -ter Art.

4. Satz. Sei $t \geq 1$. Der Rand einer E^t ist Vereinigungsmenge der Intervalle I^d ($d = 0, \dots, t-1$), welchen ein I^t von E^t zugeordnet ist.

Der Beweis verläuft analog zu dem von Satz 1. Sei b die Spitze von E^t . Sei irgendein Punkt p eines I^d gegeben, dem ein $I^t \subset E^t$ zugeordnet ist. Gesetzt, p wäre keine obere Grenze. Da auf der halboffenen Strecke $[b, p]$ nur Punkte von I^t liegen, müßte dann p einem $I^t \subset E^t$ angehören, was wegen $d \neq t$ und Satz 2 nicht sein kann.

Sei nun p obere Grenze, die keinem I^d angehört. (Es ist zu beweisen, daß p nicht existiert.)

Wegen Hilfssatz 4 liegt p in der abgeschlossenen Hülle eines I^t von E^t . Dieses I^t werde gegeben durch

$$(14) \quad x_m = \alpha_{m,m} \quad (m = s_1, \dots, s_r) \quad \max_{m=s_1, \dots, s_r} (\alpha_{m,q}) \leq x_q < \alpha_{q,q} \quad (q = v_1, \dots, v_t).$$

Da p als obere Grenze I^t selbst nicht angehören kann, andererseits I^t dem I^0 $x_m = \alpha_{m,m}$ ($m = 1, \dots, n$) zugeordnet ist, kann p im Falle $t = 1$ in der Tat nicht existieren. Im Falle $t \geq 2$ liegt er im Bereich

$$(15) \quad x_m = \alpha_{m,m} \quad (m = s_1, \dots, s_r, \sigma_1, \dots, \sigma_\tau) \quad \max_{m=s_1, \dots, s_r} (\alpha_{m,q}) \leq x_q < \alpha_{q,q} \\ (q = \omega_1, \dots, \omega_\tau),$$

wo $\varrho \geq 1$, $\tau \geq 1$, $\varrho + \tau = t$, $(\sigma_1, \dots, \sigma_\tau, \omega_1, \dots, \omega_\tau) = (v_1, \dots, v_t)$ (bis auf Permutation der Elemente).

Dieser Bereich entspricht einer Seitenfläche eines extremen Quaders im Beweis zu Satz 1. Dem extremen Quader entspricht das t -Intervall I^t .

(15) enthält das τ -Intervall

$$x_m = \alpha_{m,m} \quad (m = s_1, \dots, s_r, \sigma_1, \dots, \sigma_\varrho) \quad \max_{m=s_1, \dots, s_r, \sigma_1, \dots, \sigma_\varrho} (\alpha_{m,q}) \leq x_q < \alpha_{q,q} \\ (q = \omega_1, \dots, \omega_\tau).$$

Diesem Intervall ist I^t zugeordnet, p muß also in dem Restbereich von (15) liegen. Dieser ist im Falle $\tau \geq 2$ Vereinigungsmenge derjenigen der τ Stücke

$$x_m = \alpha_{m,m} \quad (m = s_1, \dots, s_r, \sigma_1, \dots, \sigma_\varrho) \\ \max_{m=s_1, \dots, s_r} (\alpha_{m,q}) \leq x_q < \max_{m=s_1, \dots, s_r, \sigma_1, \dots, \sigma_\varrho} (\alpha_{m,q}) \\ \max_{m=s_1, \dots, s_r} (\alpha_{m,k}) \leq x_k < \alpha_{k,k} \quad (k \neq q, s_1, \dots, s_r, \sigma_1, \dots, \sigma_\varrho),$$

($q = \omega_1, \dots, \omega_r$), die wirklich existieren. Es kann ja z. B.

$$\max_{m=s_1, \dots, s_r} (\alpha_m \omega_1) = \max_{m=s_1, \dots, s_r, v_1, \dots, v_\varrho} (\alpha_m \omega_1)$$

sein. Im Falle $\tau = 1$ wird der Restbereich, sofern er nicht leer ist, gegeben durch

$$x_m = \alpha_m m \quad (m = s_1, \dots, s_r, \sigma_1, \dots, \sigma_\varrho)$$

$$\max_{m=s_1, \dots, s_r} (\alpha_m \omega_1) \leq x_{\omega_1} < \max_{m=s_1, \dots, s_r, v_1, \dots, v_\varrho} (\alpha_m \omega_1).$$

Existiert keines der Stücke, ist also das τ -Intervall mit dem Bereich (14) identisch, so existiert p nicht und der Beweis ist fertig. Existiert wenigstens eines, so dürfen wir annehmen, daß p in dem durch $q = \omega_1$ gegebenen Stück liegt (Permutation der ω_m). Wir dürfen weiter annehmen, daß

$$\max_{m=s_1, \dots, s_r, v_1, \dots, v_\varrho} (\alpha_m \omega_1) = \alpha_{v_\varrho} \omega_1 \text{ ist (Permutation der } \sigma_m). \text{ Wir nehmen den}$$

Quader P

$$0 < x_m < \alpha_m m \quad (m \neq \omega_1) \quad 0 < x_{\omega_1} < \max_{m=s_1, \dots, s_r, \sigma_1, \dots, \sigma_\varrho} (\alpha_m \omega_1).$$

Er ist frei, da er in $(\alpha_{p,q})$ enthalten ist. Er enthält auf der $(s_1), \dots, (\sigma_{\varrho-1})$ -Seitenfläche bzw. die M -Punkte $a_{s_1}, \dots, a_{\sigma_{\varrho-1}}$, auf der $(\sigma_\varrho \omega_1)$ -Seitenfläche den M -Punkt a_{v_ϱ} . Ich bilde den Quader $(\beta_{p,q}) = d_{\sigma_\varrho} d_{v_\varrho} \dots d_{v_\tau}(P)$. Er ist n -ter Art. Es gilt

$$a_{s_1} = b_{s_1}, \dots, a_{\sigma_{\varrho-1}} = b_{\sigma_{\varrho-1}}, a_{v_\varrho} = b_{v_\varrho},$$

ferner

$$\beta_{\sigma_\varrho \sigma_\varrho} > \alpha_{\sigma_\varrho \sigma_\varrho}, \beta_{v_\varrho v_\varrho} \geq \alpha_{v_\varrho v_\varrho}, \dots, \beta_{v_\tau v_\tau} \geq \alpha_{v_\tau v_\tau}.$$

Da p in dem durch $q = \omega_1$ gegebenen Stück liegt, ist p also enthalten in dem Bereich

$$(15') \quad x_m = \beta_m m \quad (m = s_1, \dots, s_r, \sigma_1, \dots, \sigma_{\varrho-1}) \quad \max_{m=s_1, \dots, s_r} (\beta_m k) \leq x_k < \beta_k k$$

$$(k = \sigma_\varrho, \omega_1, \dots, \omega_r).$$

(15') liegt nun im Falle $\varrho > 1$ in der abgeschlossenen Hülle des Intervalls

$$(14') \quad x_m = \beta_m m \quad (m = s_1, \dots, s_r) \quad \max_{m=s_1, \dots, s_r} (\beta_m k) \leq x_k < \beta_k k$$

$$(k = v_1, \dots, v_t),$$

welches den gleichen Anfangspunkt wie (14) hat, also E^t angehört, im Falle $\varrho = 1$ in (14') selbst; im letzteren Falle ist der Beweis beendet. Im Falle $\varrho > 1$ wenden wir auf (15') dieselben Schlüsse an wie auf (15) u.s.f. und kommen nach endlich vielen Schritten zum Ziel.

5. Satz. *Liegt ein I^d im Rande einer E^t ($t \geq 1$), so auch die ganze d -Ecke, welcher I^d angehört.*

Beweis. Auf Grund von Satz 4 genügt es zu beweisen: Sind zwei I^d , K^d mit gleichem Anfangspunkt gegeben, und ist dem I^d I^t zugeordnet, so ist dem K^d ein Intervall K^t zugeordnet, welches den gleichen Anfangspunkt hat wie I^t .

Die den Intervallen I^d , K^d zugeordneten Quader seien $(\alpha_{pq}), (\beta_{pq})$. Die Koordinaten des Anfangspunktes von I^d seien $\xi_q = \max_{m=s'_1, \dots, s'_u} (\alpha_{mq})$ ($q = 1, \dots, n$), von K^d $\eta_q = \max_{m=s'_1, \dots, s'_u} (\beta_{mq})$ ($q = 1, \dots, n$), wobei $u = n - d$. Der Anfangspunkt von I^t sei $\xi'_q = \max_{m=\sigma_1, \dots, \sigma_r} (\alpha_{mq})$ ($r < u$), wobei wir $\sigma_1 = s_1, \dots, \sigma_r = s_r$ voraussetzen dürfen. Nun gilt $\xi_q = \eta_q$, also $\alpha_{s_i} = \beta_{s'_i(i)}$ ($i = 1, \dots, u$). Hierbei sind die k (i) in $(1, \dots, u)$ enthalten. Es folgt $\xi'_q = \max (\beta_{s'_i(i)q}, \dots, \beta_{s'_i(r)q})$. Da $(s'_{i(1)}, \dots, s'_{i(r)}) \subset (s'_1, \dots, s'_u)$, sind die ξ'_q die Koordinaten des Anfangspunktes eines Intervalles K^t , welches K^d zugeordnet ist.

Unter dem Symbol I^t_i verstehen wir im folgenden ein 1-Intervall, welches der x_i -Achse unseres Koordinatensystems parallel liegt.

Wir gehen jetzt zum Beweis des Hilfsatzes 8 (s. u.) über. Sei $t \geq 2$. Sei ein I^0 vorgelegt, dem der Quader (α_{pq}) zugeordnet sein möge. Die 1-Intervalle

$$(16) \quad x_m = \alpha_{mm} \quad (m \neq l) \quad \max_{m \neq l} (\alpha_{ml}) \leq x_l < \alpha_{ll}$$

sind die I^0 zugeordneten 1-Intervalle. Wir nehmen von ihnen t beliebige, $I^1_{l_1}, \dots, I^1_{l_t}$, wobei l_1, \dots, l_t paarweis verschieden unter den Zahlen $1, \dots, n$ gewählt sind. Jedem dieser Intervalle und dem Intervall I^0 ist das t -Intervall I^t

$$(17) \quad \begin{aligned} x_m &= \alpha_{mm} & (m \neq l_1, \dots, l_t), \\ \max_{m \neq l_1, \dots, l_t} (\alpha_{mq}) &\leq x_q < \alpha_{qq} & (q = l_1, \dots, l_t). \end{aligned}$$

zugeordnet. Im Rande der t -Ecke E^t , welche I^t enthält, liegen wegen Satz 3 und $t \geq 2$ die Intervalle $I^0, I^1_{l_1}, \dots, I^1_{l_t}$.

Sei nun irgendeine t -Ecke H^t gegeben, deren Rand $I^0, I^1_{l_1}, \dots, I^1_{l_t}$ enthält. Jedem dieser Intervalle muß nach Satz 3 ein t -Intervall aus H^t zugeordnet sein. Nun sind dem Intervall I_{l_1} die t -Intervalle

$$(18) \quad \begin{aligned} x_m &= \alpha_{mm} & (m \neq v_1, \dots, v_t), \\ \max_{m \neq v_1, \dots, v_t} (\alpha_{mq}) &\leq x_q < \alpha_{qq} & (q = v_1, \dots, v_t) \end{aligned}$$

zugeordnet, wo (v_1, \dots, v_t) eine vollständige Schar von Systemen $\subset (1, \dots, n)$ durchläuft, die alle l_1 enthalten und sich durch Permutation ihrer Elemente nicht ineinander überführen lassen, entsprechend dem I_{l_1} die Intervalle (18), wo (v_1, \dots, v_t) eine vollständige Schar l_2 enthaltender Systeme durchläuft usf. H^t ist also mit einer derjenigen t -Ecken identisch, welche von jeder dieser Intervallscharen (18) mindestens ein Intervall enthalten. Um diese t -Ecken zu bestimmen, müssen wir die Mengen von (nicht notwendig paarweise verschiedenen) Intervallen mit gleichem Anfangspunkt aufstellen, die aus jeder Schar (18) ein Element enthalten. Damit das l_1 enthaltende System (v_1, \dots, v_t) und das l_2 enthaltende System (v'_1, \dots, v'_t) Intervalle (18) mit gleichem Anfangspunkt liefern, muß l_2 in (v_1, \dots, v_t) und l_1 in (v'_1, \dots, v'_t) enthalten sein. Damit das l_1, l_2 enthaltende System (v_1, \dots, v_t) und das l_3 enthaltende System (v'_1, \dots, v'_t) Intervalle (18) mit gleichem Anfangspunkt liefern, muß l_3 in (v_1, \dots, v_t) und l_1, l_2 in (v'_1, \dots, v'_t) enthalten sein usf. Es stellt sich heraus, daß nur eine Menge von Intervallen mit gleichem Anfangspunkt existiert, für welche jede Schar (18) ein Element liefert, nämlich diejenige Menge, die das Intervall (17) t -mal enthält. Also ist H^t eindeutig bestimmt.

Ist eine t -Ecke H^t gegeben, in deren Rand I^0 liegt, so muß sie nach Satz 4 ein I^t enthalten, welches I^0 zugeordnet ist. Dieses I^t ist aber $t-1$ -Intervallen zugeordnet, die ihrerseits I^0 zugeordnet sind, so daß wir durch die Systeme (l_1, \dots, l_t) auch alle t -Ecken erhalten, in deren Rand I^0 liegt.

Ist ein System (l'_1, \dots, l'_t) gegeben, welches sich durch Permutation seiner Elemente nicht in (l_1, \dots, l_t) überführen läßt, bilden wir analog (17) mit diesen Systemen die t -Intervalle I'^t, I^t , so haben diese Intervalle verschiedene Anfangspunkte, liegen also in verschiedenen t -Ecken. Insgesamt folgt der

8. Hilfssatz. Sei $2 \leq t \leq n-1$. Ist ein I^0 vorgegeben und irgendein in $(1, \dots, n)$ enthaltenes System paarweis verschiedener Zahlen (l_1, \dots, l_t) , so liegen I^0 und die I^0 zugeordneten Intervalle $I^1_{l_1}, \dots, I^t_{l_t}$ im Rande einer eindeutig bestimmten E^t .

Jede E^t , in deren Rand I^0 liegt, kann durch Wahl eines Systems (l_1, \dots, l_t) erhalten werden.

Sind zwei Systeme $(l_1, \dots, l_t), (l'_1, \dots, l'_t)$ vorgegeben, die sich durch Permutation ihrer Elemente nicht ineinander überführen lassen, so sind auch die t -Ecken verschieden, in deren Rand $I^0, I^1_{l_1}, \dots, I^t_{l_t}$ bzw. $I^0, I^1_{l'_1}, \dots, I^t_{l'_t}$ liegen.

Es folgt der

9. Hilfssatz. Sei $2 \leq u < t \leq n-1$. Ist E^t die gemäß Hilfssatz 8 durch ein I^0 und irgendwelche zugeordneten Intervalle $I^1_{l_1}, \dots, I^t_{l_t}$ bestimmte t -Ecke (l_1, \dots, l_t) paarweis verschiedene unter den Zahlen $1, \dots, n$, ist (h_1, \dots, h_u) ein in (l_1, \dots, l_t) enthaltenes System paarweis verschiedener Zahlen, F^u die gemäß Hilfssatz 8 durch $I^0, I^1_{h_1}, \dots, I^u_{h_u}$ bestimmte u -Ecke, so liegt F^u im Rand von E^t .

Beweis. Sei wieder I^0 Eckpunkt des Quaders (α_{pq}) n -ter Art. In E^t liegt dann, wie wir sahen, das Intervall (17). In F^u liegt entsprechend das Intervall

$$(17') \quad \begin{aligned} x_m &= \alpha_{mm} & (m \neq h_1, \dots, h_u), \\ \max_{m \neq h_1, \dots, h_u} (\alpha_{mq}) &\leq x_q < \alpha_{qq} & (q = h_1, \dots, h_u). \end{aligned}$$

Wegen $(h_1, \dots, h_u) \subset (l_1, \dots, l_t)$ und $u < t$ ist aber (17) dem Intervall (17') zugeordnet, also (17') nach Satz 3 im Rand von E^t gelegen, also wegen Satz 4 die ganze F^u .

10. Hilfssatz. Sei $1 \leq t \leq n-1$. Seien die 0-Ecken des Randes einer E^t $p^k = (\pi_1^k, \dots, \pi_n^k)$. Die Spitze von E^t wird dann gegeben durch $\tau_m = \min_k (\pi_m^k)$ ($m = 1, \dots, n$), d. h. E^t ist durch die 0-Ecken ihres Randes eindeutig bestimmt.

Beweis. Sei (τ_{pq}) der Quader r -ter Art, dessen Eckpunkt Spitze von E^t ist. Einem I^0 p^k des Randes von E^t ist ein t -Intervall aus E^t

$$\begin{cases} x_m = \tau_{mm} = \pi_m^k & (m = s_1, \dots, s_r), \\ \tau_{mm} \leq x_m < \pi_m^k & (m \neq s_1, \dots, s_r) \end{cases}$$

zugeordnet. Demnach gilt

$$(19) \quad \tau_{mm} \leq \pi_m^k \text{ für alle } m \text{ und } k.$$

Andererseits liegt in jedem Bereich

$$x_m = \tau_{mm} \quad (m = s_1, \dots, s_r) \quad \tau_{mm} \leq x_m \quad (m \neq s_1, \dots, s_r),$$

wobei (s_1, \dots, s_r) ein System von Zahlen ist, das aus jeder Serie des Typus von (τ_{pq}) genau ein Element enthält, nach Hilfssatz 2 und Satz 4 mindestens ein 0-Intervall $p^{s_1, \dots, s_r} = (\pi_1^{s_1, \dots, s_r}, \dots, \pi_n^{s_1, \dots, s_r})$ des Randes von E^t . Bedenken wir, daß in den Serien des Typus von (τ_{pq}) jede der Zahlen $(1, \dots, n)$ vorkommt, so kann ich zu einer vorgegebenen dieser Zahlen, etwa l , ein solches System (s_1, \dots, s_r) wählen, das l enthält. Alsdann ist $\pi_l^{s_1, \dots, s_r} = \tau_{ll}$. Daraus folgt mit (19) $\tau_{mm} = \min_k \pi_m^k$ ($m = 1, \dots, n$), w. z. b. w.

11. Hilfssatz. Sei $1 \leq d < t < n-1$. Liegen die 0-Ecken des Randes einer E^d im Rand von E^t , so liegt die ganze E^d im Rand von E^t .

Beweis. Seien p^1, \dots, p^h die 0-Ecken des Randes von E^d , $p^1, \dots, p^h, \dots, p^w$ die des Randes von E^t , sei $p^k = (\pi_1^k, \dots, \pi_n^k)$ ($k = 1, \dots, w$).

Ich nehme ein p^k mit $1 \leq k \leq h$, der zugeordnete Quader sei (π_{pq}) . p^k ist ein t -Intervall I^t von E^t zugeordnet, ein d -Intervall I^d von E^d . I^t werde gegeben durch

$$x_m = \pi_{mm} \quad (m = s_1, \dots, s_r) \quad \max_{m = s_1, \dots, s_r} (\pi_{mq}) \leq x_q < \pi_{qq} \quad (q = v_1, \dots, v_i).$$

1. Angenommen, I^t ist I^d zugeordnet, so liegt nach Satz 4 I^d im Rande von E^t und nach Satz 5 auch E^d .

2. Ist I^t I^d nicht zugeordnet, so wird I^d gegeben durch

$$x_m = \pi_{m m} \quad (m = s'_1, \dots, s'_g) \quad \max_{m=s'_1, \dots, s'_g} (\pi_{m q}) \leq x_q < \pi_{q q} \quad (q = v'_1, \dots, v'_t),$$

wobei $g = n - d$ und wenigstens eine Zahl v'_u mit einem s'_g übereinstimmt. Seien $\mathbf{x}^d = (\xi_1^d, \dots, \xi_n^d)$, $\mathbf{x}^t = (\xi_1^t, \dots, \xi_n^t)$ bzw. die Anfangspunkte von I^d , I^t . Dann gilt

$$\xi_{v'_u}^d = \max (\pi_{s'_1 v'_u}, \dots, \pi_{s'_g v'_u}) < \pi_{v'_u v'_u} = \pi_{s'_g s'_g} = \xi_{s'_g}^t,$$

$$\xi_{v'_u}^d < \xi_{s'_g}^t.$$

Nun ist aber nach Hilfssatz 10 $\xi_m^t = \min_{k=1, \dots, h} (\pi_m^k)$, $\xi_m^d = \min_{k=1, \dots, w} (\pi_m^k)$ ($m = 1, \dots, n$), also

$$\xi_m^d \geq \xi_m^t \quad (m = 1, \dots, n),$$

so daß 2. ausscheidet und der Satz bewiesen ist.

5. Satz. Jede 1-Ecke enthält genau 2 1-Intervalle I_a^1 , I_b^1 mit $a \neq b$.

Beweis. Der Anfangspunkt b einer 1-Ecke ist nach Hilfssatz 1 Eckpunkt eines Quaders $(n-1)$ -ter Art. Der Typus eines solchen Quaders enthält notwendig $(n-1)$ Serien, ist also von der Gestalt

$$(l_1 l_2) (l_3) \dots (l_n).$$

Nach Hilfssatz 2 gibt es ein 1-Intervall mit b als Anfangspunkt nur in der Parallelen zur x_1 -Achse durch b und der Parallelen zur x_2 -Achse durch b . Wegen Satz 2 gibt es in jeder dieser Parallelen nur ein I^1 mit b als Anfangspunkt. Da $l_1 \neq l_2$, ist der Satz bewiesen.

Auf dem Rande einer E^1 liegen nach Satz 4 genau die 0-Intervalle, denen die 1-Intervalle von E^1 zugeordnet sind, d. h. genau zwei. Jeder 1-Ecke ordnen wir nun einen Prozeß zu folgendermaßen: Liegen die beiden 0-Ecken P^0 , S^0 im Rande einer E^1 , die aus den Intervallen I_1^1 , I_k^1 besteht, und ist I_1^1 der P^0 , I_k^1 der S^0 zugeordnet, so sagen wir, S^0 geht aus P^0 durch den Prozeß Ω_{1k} hervor,

$$S^0 = \Omega_{1k}(P^0).$$

Definitionsgemäß gilt dann

$$P^0 = \Omega_{k1}(S^0).$$

Wir sagen auch, Ω_{k1} ist auf S^0 anwendbar, Ω_{1k} auf P^0 .

S^0 , P^0 können nicht im Rande zweier verschiedener 1-Ecken liegen wegen Hilfssatz 10, so daß aus $S^0 = \Omega_{cd}(P^0)$ $c = l$, $d = k$ folgt.

Ist ein $P^0 = I^0$ gegeben, und eine beliebige der Zahlen $1, \dots, n$, etwa l , so gibt es genau eine E^1 mit I^0 im Rande, die das I^0 zugeordnete I_l^1 enthält.

Infolgedessen ist auch ein Prozeß Ω_{ik} auf I^0 anwendbar, wobei $k = k(l, I^0)$ und $\Omega_{ik(l, I^0)}(I^0) = S^0$ eindeutig durch l , I^0 bestimmt und $k(l, I^0) \neq l$ ist.

Die Möglichkeit der beliebigen Wahl von l sowie die Eindeutigkeit von Ω_{ik} und k können auch den unten folgenden Überlegungen entnommen werden.

Ist $l \neq l'$, so sind die 1-Ecken mit I^0 , $\Omega_{ik(l, I^0)}(I^0)$ bzw. I^0 , $\Omega_{i'k'(l', I^0)}(I^0)$ auf dem Rande verschieden.

Wir definieren jetzt Operationen für die extremen Quader n -ter Art. Sei $P = (\pi_{pg})$ n -ter Art. Sei l eine der Zahlen $1, \dots, n$. Ich komprimiere P in Richtung der negativen x_l -Achse solange, bis ich mit der abgeschlossenen (l) -Seitenfläche auf den ersten von $p_l = (\pi_{l1}, \dots, \pi_{ln})$ verschiedenen M -Punkt der Begrenzung von P stoße. Sei dies der M -Punkt $p_{f(l), p_l}$. Der so erhaltene Quader heiße Q . Ich dilatiere jetzt Q in Richtung der positiven x_l -Achse, bis ich erstmalig auf einen M -Punkt stoße. Der so erhaltene Quader sei $S = (\sigma_{pg})$. Wir sagen, S sei aus P durch die Operation Ω_{lf} hervorgegangen, und schreiben $S = \Omega_{lf}(P)$. Ω_{lf} heißt auf P anwendbar.

Q ist extrem $(n-1)$ -ter Art, S extrem n -ter Art, wie gleich bewiesen wird. Es folgt:

1. Ist P^0 der Eckpunkt von P , S^0 der von S , so gilt

$$S^0 = \Omega_{lf}(P^0).$$

Dazu ist zu beweisen, daß P^0 , S^0 den Rand einer 1-Ecke bilden, daß diese 1-Ecke zwei Intervalle I_l^1 , $I_{l'}^1$ enthält und I_l^1 dem P^0 , $I_{l'}^1$ dem S^0 zugeordnet ist.

Der Quader Q wird gegeben durch

$$0 < x_m < \pi_{m m} \quad (m \neq l) \quad 0 < x_l < \max_{m \neq l} (\pi_{m l}) = \pi_{f l}.$$

Er enthält auf seiner (m) -Seitenfläche $(m \neq f, l)$ den M -Punkt $p_m = (\pi_{m1}, \dots, \pi_{m m})$, auf seiner (lf) -Seitenfläche p_f . Der Quader $S = d_f(Q)$ ist demnach n -ter Art; setzen wir $j_m = (\sigma_{m1}, \dots, \sigma_{m m})$ $(m = 1, \dots, n)$, so gilt $j_m = p_m$ $(m \neq l, f)$; $j_l = p_f$. Außerdem ist $\sigma_{ff} > \pi_{ff}$.

Die gesuchte 1-Ecke, deren Randpunkte P^0 , S^0 sind, wird dann geliefert von den beiden 1-Intervallen

$$\begin{aligned} I_l^1: \quad x_m &= \pi_{m m} \quad (m \neq l) & \max_{m \neq l} (\pi_{m l}) &\leq x_l < \pi_{l l}, \\ I_{l'}^1: \quad x_m &= \sigma_{m m} \quad (m \neq f) & \max_{m \neq f} (\sigma_{m f}) &\leq x_f < \sigma_{f f}. \end{aligned}$$

Beide Intervalle haben nach den hergeleiteten Beziehungen gleichen Anfangspunkt und I_l^1 ist P^0 , $I_{l'}^1$ S^0 zugeordnet, w. z. b. w.

2. Ist für zwei 0-Ecken P^0 , S^0 $S^0 = \Omega_{ik}(P^0)$, so gilt für die Quader P , S mit den Eckpunkten P^0 , S^0

$$S = \Omega_{ik}(P).$$

Beweis. Die 1-Ecke mit den Randpunkten P^0, S^0 enthält ein I_l^1 , welches P^0 zugeordnet ist und ein I_k^1 , welches S^0 zugeordnet ist:

$$I_l^1: x_m = \pi_{m m} \quad (m \neq l) \quad \max_{m \neq l} (\pi_{m l}) \leq x_l < \pi_{l l},$$

$$I_k^1: x_m = \sigma_{m m} \quad (m \neq k) \quad \max_{m \neq k} (\sigma_{m k}) \leq x_k < \sigma_{k k}.$$

Wegen der Gleichheit der Anfangspunkte haben wir für $m \neq l, k$ $\sigma_{m m} = \pi_{m m}$, ferner $\max_{m \neq l} (\pi_{m l}) = \sigma_{l l}$, $\max_{m \neq k} (\sigma_{m k}) = \pi_{k k}$. Daraus folgt insbesondere $\max_{m \neq l} (\pi_{m l}) = \pi_{k l}$. Wir stoßen also bei Kompression des Quaders P in Richtung der negativen x_l -Achse erstmalig auf den M -Punkt p_k der (k) -Seitenfläche von P . Der Quader Q wird gegeben durch

$$0 < x_m < \pi_{m m} \quad (m \neq l) \quad 0 < x_l < \pi_{k l}.$$

Er kann wegen der oben hergeleiteten Beziehungen auch in der Form

$$0 < x_m < \sigma_{m m} \quad (m \neq k) \quad 0 < x_k < \max_{m \neq k} (\sigma_{m k})$$

geschrieben werden. Dilatieren wir diesen Quader in Richtung der positiven x_k -Achse, so stoßen wir erstmalig auf den Gitterpunkt j_k der (k) -Seitenfläche von S , da S extrem n -ter Art ist. Damit ist der Beweis von 2. erbracht.

Das für die extremen Quader n -ter Art definierte Iterationsverfahren der Ω -Prozesse, welche extreme Quader wieder in solche überführen, ist eine genaue Verallgemeinerung des von Minkowski für dreidimensionale Gitter definierten Verfahrens der Bildung von Nachbarn. Ordnen wir unserem Verfahren einen Graphen Γ^1 zu, der als Eckpunkte die Quader n -ter Art, als Strecken die Operationen $\Omega_{l k}$ enthält, so ist nach dem eben Bewiesenen dieser Graph in $B(V)$ realisiert durch die Menge der 0- und 1-Ecken.

6. Satz. Sei $1 \leq t \leq n-1$. Ist ein I^0 gegeben und irgendein System von paarweis verschiedenen Ziffern $(l_1, \dots, l_t) \subset (1, \dots, n)$, so gibt es genau eine t -Ecke mit $I^0, \Omega_{m k(m, I^0)}(I^0)$ ($m = l_1, \dots, l_t$) auf dem Rande.

Ist eine t -Ecke E^t gegeben mit I^0 als Randpunkt, so gibt es unter den Ziffern $1, \dots, n$ genau t, l_1, \dots, l_t , für welche $\Omega_{m k(m, I^0)}(I^0)$ ($m = l_1, \dots, l_t$) im Rande von E^t liegen.

Sind zwei Systeme $(l_1, \dots, l_t), (l'_1, \dots, l'_t)$ vorgegeben, die sich durch Permutation ihrer Elemente nicht ineinander überführen lassen, so sind auch die t -Ecken verschieden, in deren Rand $I^0, \Omega_{m k(m, I^0)}(I^0)$ ($m = l_1, \dots, l_t$) bzw. $I^0, \Omega_{m k(m, I^0)}(I^0)$ ($m = l'_1, \dots, l'_t$) liegen.

Beweis. Für $t=1$ ist der Satz klar bzw. auf S. 25 schon bewiesen. Nun für $t \geq 2$ zur 1. Beh.: Nach Hilfssatz 8 gibt es eine t -Ecke E^t , in deren Rand I^0 und die I^0 zugeordneten Intervalle $I_{l_1}^1, \dots, I_{l_t}^1$ liegen. Nach Satz 5 liegen

dann auch die 1-Ecken im Rande von E^t , welche bzw. $I_{l_1}^1, \dots, I_{l_t}^1$ enthalten, damit auch ihre Randpunkte im Rand von E^t . Also sind die $\Omega_{m \ k \ (m, I^0)}(I^0)$ im Rand von E^t enthalten.

Ist nun eine E^t gegeben mit I^0 und $\Omega_{m \ k \ (m, I^0)}(I^0)$ auf dem Rande, so enthält nach Hilfssatz 11 der Rand von E^t die t 1-Ecken, welche I^0 mit den $\Omega_{m \ k \ (m, I^0)}(I^0)$ verbinden, damit die I^0 zugeordneten 1-Intervalle I_1^1, \dots, I_t^1 , also ist nach Hilfssatz 8 E^t eindeutig bestimmt.

2. Beh.: Nach Hilfssatz 8 gibt es ein System $(l_1, \dots, l_t) \subset (1, \dots, n)$ von paarweis verschiedenen Zahlen, so daß mit I^0 auch die I^0 zugeordneten $I_{l_1}^1, \dots, I_{l_t}^1$ im Rand von E^t liegen. Dann liegen nach Satz 5 auch die $I_{l_1}^1, \dots, I_{l_t}^1$ enthaltenden 1-Ecken im Rande von E^t , also auch ihre Randpunkte, also auch die $\Omega_{m \ k \ (m, I^0)}(I^0)$ ($m = l_1, \dots, l_t$).

3. Beh.: Nach Hilfssatz 11 liegen im Rand der einen t -Ecke I^0 und die zugeordneten $I_{l_1}^1, \dots, I_{l_t}^1$ im Rande der anderen I^0 und die zugeordneten $I_{l_1}^1, \dots, I_{l_t}^1$. Mit der 3. Beh. von Hilfssatz 8 folgt also die unseres Satzes.

7. Satz. Sei $1 \leq u < t \leq n - 1$. Ist eine t -Ecke E^t gegeben mit I^0 , $\Omega_{m \ k \ (m, I^0)}(I^0)$ ($m = l_1, \dots, l_t$) auf dem Rande, und irgendein System paarweis verschiedener Zahlen $(h_1, \dots, h_u) \subset (l_1, \dots, l_t)$, so gibt es eine u -Ecke E^u im Rand von E^t , in deren Rand I^0 , $\Omega_{m \ k \ (m, I^0)}(I^0)$ ($m = h_1, \dots, h_u$) liegen.

Beweis. Sei $u \geq 2$. Nach Hilfssatz 11 liegen

I^0 und die zugeordneten I_m^1 ($m = l_1, \dots, l_t$) im Rande von E^t ,

I^0 und die zugeordneten I_m^1 ($m = h_1, \dots, h_u$) im Rande einer nach Satz 6 eindeutig bestimmten E^u .

Also folgt die Beh. mit Berücksichtigung von Hilfssatz 9. Für $u = 1$ folgt die Beh. sofort aus Hilfssatz 11.

8. Satz. Seien b, c ganze Zahlen mit $0 \leq b < c \leq n - 1$:

Eine E^b liegt im Rande von $\binom{n-b}{n-c} E^c$.

Seien b, c, d ganze Zahlen mit $0 \leq b < c < d \leq n - 1$:

Liegt E^b im Rand von E^d , so gibt es im Rand von E^d $\binom{d-b}{d-c} E^c$, in deren Rand E^b liegt.

Beweis. 1. Beh.: folgt für $b = 0$ $c \geq 1$ sofort aus Satz 6. Sei also $b \geq 1$, $c \geq 2$.

Wir nehmen ein I^0 im Rand von E^b . Es gibt nach Satz 6 b Ziffern (l_1, \dots, l_b) , für die $\Omega_{m \ k \ (m, I^0)}(I^0)$ ($m = l_1, \dots, l_b$) im Rand von E^b liegen. Es gibt genau c Ziffern (d_1, \dots, d_c) , für welche $\Omega_{m \ k \ (m, I^0)}(I^0)$ ($m = d_1, \dots, d_c$) im Rand von E^c liegen. Also gilt $(l_1, \dots, l_b) \subset (d_1, \dots, d_c)$.

Ist irgendein System $(d_1, \dots, d_c) \supset (l_1, \dots, l_b)$ vorgelegt, so gibt es nach Satz 6 eine E^c mit $\Omega_{m \ k \ (m, I^0)}(I^0)$ ($m = d_1, \dots, d_c$) auf dem Rande, in deren

Rand nach Satz 7 E^b liegt. Für verschiedene Systeme (d_1, \dots, d_s) erhalten wir nach Satz 6 verschiedene E^c . Also ist die Anzahl der fraglichen $E^c \binom{n-b}{n-c}$, w. z. b. w.

Der Beweis der 2. Beh. folgt entsprechend.

9. Satz. Sei eine t -Ecke E^t gegeben, $t \geq 2$. Das 0-Intervall p liege im Rand von E^t , und l_1, \dots, l_t seien die Ziffern, für welche $p_m = \Omega_{l_m k(l_m, p)}(p)$ ($m = 1, \dots, t$) im Rand von E^t liegen. Sei ein p_i gegeben ($i = 1, \dots, t$). Die Ziffern l'_1, \dots, l'_t , für welche $\Omega_{l'_m k(l'_m, p_i)}(p_i)$ ($m = 1, \dots, t$) im Rand von E^t liegen, werden dann gegeben durch

$$\left. \begin{aligned} l'_m &= l_m & \text{für } m \neq i \\ l'_i &= k(l_i, p) \end{aligned} \right\}, \text{ wenn } k(l_i, p) \notin l_1, \dots, l_m,$$

$$l'_m = l_m \quad \text{für alle } m, \quad \text{wenn } k(l_i, p) \subset l_1, \dots, l_m.$$

Beweis. p sei Eckpunkt des Quaders $(\varrho_{p,p})$ n -ter Art. Nach Hilfssatz 11 liegen die p zugeordneten Intervalle $I^1_{l_1}, \dots, I^1_{l_t}$ im Rand von E^t . Nach dem Beweis zu Hilfssatz 8 ist das Intervall

$$(17'') \quad x_g = \varrho_{g,g} \quad (g \neq l_1, \dots, l_t) \quad \max_{g \neq l_1, \dots, l_t} (\varrho_{g,h}) \leq x_h < \varrho_{h,h} \quad (h = l_1, \dots, l_t)$$

das p in E^t zugeordnete t -Intervall. Wir dürfen ohne Beschränkung der Allgemeinheit $i = 1$ annehmen.

p_1 sei Eckpunkt des Quaders $(\sigma_{p,p})$ n -ter Art. p_1, p sind die Randpunkte einer 1-Ecke F^1 , die aus einem $I^1_{l_1}$ und einem $I^1_{k(l_1, p)}$ besteht, wobei das erste Intervall p , das zweite p_1 zugeordnet ist. Die beiden Intervalle werden also gegeben durch

$$x_g = \varrho_{g,g} \quad (g \neq l_1) \quad \max_{g \neq l_1} (\varrho_{g,l_1}) \leq x_{l_1} < \varrho_{l_1,l_1},$$

und

$$x_g = \sigma_{g,g} \quad (g \neq k) \quad \max_{g \neq k} (\sigma_{g,k}) \leq x_k < \sigma_{k,k},$$

wobei $k = k(l_1, p)$ gesetzt ist. Es folgt wegen der Gleichheit der Anfangspunkte

$$r_g = l_g \quad (g \neq l_1, k), \quad \max_{g \neq l_1} (\varrho_{g,l_1}) = \sigma_{l_1,l_1}, \quad \max_{g \neq k} (\sigma_{g,k}) = \varrho_{k,k}, \quad \text{also } r_k = l_1.$$

1. k sei in (l_1, \dots, l_t) enthalten.

Dann ist $\max_{g \neq l_1, \dots, l_t} (\varrho_{g,h}) = \max_{g \neq l_1, \dots, l_t} (\sigma_{g,h})$ ($h = 1, \dots, n$), d. h. das t -Intervall $x_g = \sigma_{g,g}$ ($g \neq l_1, \dots, l_t$) $\max_{g \neq l_1, \dots, l_t} (\sigma_{g,h}) \leq x_h < \sigma_{h,h}$ ($h = l_1, \dots, l_t$) gehört zu E^t . Es ist den t Intervallen

$$(20) \quad x_g = \sigma_{g,g} \quad (g \neq l), \quad \max_{g \neq l} (\sigma_{g,l}) \leq x_l < \sigma_{l,l}$$

($l = l_1, \dots, l_t$) zugeordnet, die also nach Satz 4 im Rande von E^t liegen. Damit liegen auch die diese Intervalle enthaltenden 1-Ecken im Rand von E^t (Satz 5), damit auch ihre Randpunkte. Da die Intervalle (20) sämtlich p_1 zugeordnet sind, folgt die Beh. im Falle 1.

2. k sei von l_1, \dots, l_t verschieden.

Dann ist $\max_{g \neq l_1, \dots, l_t} (\varrho_{g,h}) = \max_{g \neq l_2, \dots, l_t, k} (\sigma_{g,h})$ ($h = 1, \dots, n$) wegen

$$\begin{aligned} \max_{g \neq l_1, \dots, l_t} (\varrho_{g,h}) &= \max [\varrho_{k,h}, \max_{g \neq l_1, \dots, l_t, k} (\varrho_{g,h})] = \\ &= \max [\sigma_{l_1,h}, \max_{g \neq l_1, \dots, l_t, k} (\sigma_{g,h})] = \max_{g \neq l_2, \dots, l_t, k} (\sigma_{g,h}), \end{aligned}$$

d. h. das t -Intervall

$$x_g = \sigma_{g,g} \quad (g \neq l_2, \dots, l_t, k) \quad \max_{g \neq l_2, \dots, l_t, k} (\sigma_{g,h}) \leq x_h < \sigma_{h,h} \quad (h = l_2, \dots, l_t, k)$$

gehört zu E^t . Jetzt führen die gleichen Schlüsse wie unter 1. zur Behauptung des Satzes.

10. Satz. Die auf dem Rande einer E^t ($t \geq 2$) liegenden 0- und 1-Ecken bilden einen zusammenhängenden Streckenkomplex.

Beweis. Die Beh. ist offenbar richtig für $t = 2$, da der Rand von E^2 in diesem Fall eine topologische Kreislinie ist, die in 0- und 1-Ecken zerfällt.

Angenommen, die Beh. sei richtig für $t = t_0$. Sei eine E^{t_0+1} gegeben, seien p, q beliebige 0-Ecken ihres Randes. Dieser Rand wird überdeckt von den abgeschlossenen Hüllen der t_0 -Ecken, die ihm zugehören. Ich verbinde p, q durch eine orientierte stetige Kurve k mit dem Anfangspunkt p , dem Endpunkt q .

Da die Ecken topologische Kugeln sind, kann ich k so wählen, daß mit zwei Punkten r, s , die der gleichen abgeschlossenen t_0 -Ecke des Randes angehören, auch das ganze zwischen r und s liegende Stück von k zu dieser abgeschlossenen t_0 -Ecke gehört.

Es gibt eine t_0 -Ecke $E'_{(0)}$, deren abgeschlossene Hülle $p = p_0$ und eine ganze Nachbarschaft von p in k enthält. Es gibt einen Punkt $p_1 \neq p_0$ auf k , für welchen alle zwischen p_0 und p_1 liegenden Punkte von k zu $E'_{(0)}$ gehören, aber kein auf p_1 folgender.

Es gibt weiter eine t_0 -Ecke $E'_{(1)}$, deren abgeschlossene Hülle p_1 und eine ganze Nachbarschaft von p_1 in dem auf p_1 folgenden Stück von k enthält. Ich bestimme den Punkt $p_2 \neq p_1$ auf k , für welchen alle zwischen p_1 und p_2 liegenden Punkte von k zur abgeschlossenen Hülle von $E'_{(1)}$ gehören usw.

Wegen der Voraussetzungen über k gelange ich nach endlich vielen Schritten zum Punkt $p_r = q$ und breche die Definition der $E'_{(m)}$, p_m ab.

Ist $f = 1$, so liegen p, q auf dem Rande der gleichen t_0 -Ecke, sind also nach Voraussetzung durch einen auf dem Rande von $E_{(0)}^{t_0}$, also auf dem Rande von $E_{(0)}^{t_0+1}$ liegenden Streckenzug von 0- und 1-Ecken verbindbar.

Sei $f \geq 2$. Ein Punkt p_m ($m = 1, \dots, f-1$) gehört zum Rande von $E_{(m-1)}^{t_0}$ und zum Rande von $E_{(m)}^{t_0}$, also zu einer F^P , deren abgeschlossene Hülle ganz dem Rande von $E_{(m-1)}^{t_0}$ und $E_{(m)}^{t_0}$ angehört. Ich wähle eine 0-Ecke r_m auf dem Rand von F^P . Sie gehört dem Rand von $E_{(m)}^{t_0}$ und $E_{(m-1)}^{t_0}$ an. Ich setze $p_0 = r_0, p_f = r_f$. Die aufeinanderfolgenden Punkte r_m, r_{m+1} liegen im Rande der gleichen t_0 -Ecke $E_{(m)}^{t_0}$, sind also nach Induktionsvoraussetzung durch einen Streckenzug von 0- und 1-Ecken im Rande von $E_{(m)}^{t_0}$, also auch im Rande von $E_{(0)}^{t_0+1}$ verbindbar, damit auch p und q .

Die Beweise der folgenden Sätze, die für $n = 3$ schon in Z 1 gebracht wurden, können ohne Schwierigkeit auf die n -te Dimension verallgemeinert werden.

Es gilt (vgl. Z 1, 2. Satz) der

11. Satz. *Ordnen wir jedem Punkt von $B(V)$ denjenigen Punkt der Fläche $x_1 \cdot x_2 \cdot \dots \cdot x_n = 1, x_1 > 0 \dots x_n > 0$ zu, der mit ihm auf derselben Geraden durch den Nullpunkt liegt, so ist diese Abbildung topologisch.*
und der (vgl. Z 1, 13. Satz)

12. Satz. *Bei der Abbildung*

$$\xi_m = \ln \frac{x_m}{\sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n}} \quad (m = 1, \dots, n)$$

des O_n auf die Hyperebene $\xi_1 + \dots + \xi_n = 0$ werden

zu einer festen Koordinatenachse parallele gleichgerichtete Vektoren auf parallele gleichgerichtete Vektoren transformiert,

zwei verschiedenen Koordinatenachsen parallele Vektoren, die beide in Richtung der positiven oder beide in Richtung der negativen Achsen weisen, auf zwei Vektoren, die den durch

$$\cos \varphi = \frac{1}{n-1}, \quad \frac{\pi}{2} < \varphi < \pi$$

bestimmten Winkel einschließen.

13. Satz. *Der aus allen 0- und 1-Ecken bestehende Streckenkomplex ist zusammenhängend.*

Der Beweis dieses Satzes stützt sich erstens darauf, daß jeder Bereich $0 < x_m \leq c_m$ ($m = 1, \dots, n, c_m$ beliebig positiv) nur mit endlich vielen n -Ecken Punkte gemein hat, zweitens darauf, daß $B(V)$ nach Satz 11 zusammenhängend ist, im übrigen verläuft er entsprechend dem Beweis von Satz 10.

(Eingegangen am 30. 4. 1940.)

Von den Verteilungen auf Booleschen Ringen.

Von

Bedřich Pospíšil in Bränn.

Es handelt sich um einige Ergänzungen zu meiner Abhandlung „Über die meßbaren Funktionen“, woraus ich die Terminologie und Bezeichnungen übernehme¹⁾. Wir betrachten einen festen Booleschen Ring A mit Einselement e .

Folgende Bemerkung wird sich von Nutzen zeigen. Bezeichnet man mit i mit etwaigen Indizes endliche Vereinigungen von Intervallen und ist $\varphi(i)$ gleich der Summe $\Sigma \varphi(j_n)$, wo j_n endlich viele zu je zwei fremde Intervalle mit $i = \Sigma j_n$ sind, so sieht man erstens, daß $\varphi(i)$ von der Aufspaltung von i unabhängig ist und zweitens

$$(\alpha) \quad \varphi(i_1 + i_2) = \varphi(i_1) + \varphi(i_2)$$

und

$$(\beta) \quad \varphi(i_1 i_2) = \varphi(i_1) \varphi(i_2)$$

gilt. Also kann man die Verteilungen auch so erklären, daß ihr Argument den Mengenkörper aller i durchläuft.

In der Tat kann man (α) und (β) durch folgende Bedingungen ersetzen²⁾:

$$(\alpha') \quad \varphi(i_1 \vee i_2) = \varphi(i_1) \vee \varphi(i_2) \quad \text{und} \quad \varphi(i') = \varphi(i)'$$

Denn erstens ist $(\alpha)(\beta) \rightarrow (\alpha')(\beta')$ naheliegend. Umgekehrt folgt (β) aus $(\alpha')(\beta')$ und $(ab)' = a' \vee b'$; (α) ist dann eine Folge von $(\alpha')(\beta)$ und $a + b = (a \vee b) \cdot (ab)'$. Man hat also $(\alpha')(\beta')$ zu beweisen. Aus der Verteilungserklärung folgt nun, daß für $j = j_1 \vee j_2 \vee \dots \vee j_N$ auch $\varphi(j) = \varphi(j_1) \vee \dots \vee \varphi(j_N)$ gilt. Also ist $\varphi(i)$ die Vereinigung beliebiger endlich vieler $\varphi(j_N)$, wenn nur j die Vereinigung der j_n ist, woraus ja (α') folgt. Ferner gilt ja $e = \varphi(i + i') = \varphi(i) + \varphi(i')$, d. h. (β') .

¹⁾ Math. Ann. 117 (1940), S. 327ff. Diese Abhandlung wird mit MF. bezeichnet. Insbesondere bleiben alle Erklärungen und Verabredungen der Einleitung von MF. (l. c., S. 327f.) bestehen. Statt $a = ab$ schreibe ich $a \subset b$.

²⁾ $a \vee b = a + ab + b$, $a_1 \vee a_2 \vee \dots \vee a_n \vee a_{n+1} = (a_1 \vee \dots \vee a_n) \vee a_{n+1}$. Diese Operation, sog. Vereinigung, ist kommutativ und assoziativ. a' ist immer das Komplement von a : $a + a'$ ist das Einselement. Für Mengen stimmt es mit der üblichen Terminologie überein. Vgl. M. H. Stone, Trans. Amer. Math. Soc. 40 (1936), S. 37ff. Mit St. 63 (z. B.) wird Theorem 63 dieser Abhandlung bezeichnet.

I.

Stetige Homomorphismen.

$|\varphi| \leq \varepsilon$ ($\varepsilon > 0$) besagt: $\varphi(j) = e$, wo j das Intervall aller x mit $|x| \leq \varepsilon$ ist. φ ist beschränkt, falls $|\varphi| \leq \varepsilon$ für ein $\varepsilon > 0$ gilt. Ich schreibe $\varphi_n \rightarrow 0$, falls es eine Nullfolge $\varepsilon_n > 0$ mit $|\varphi_n| \leq \varepsilon_n$ gibt. Eine (reelle) Funktion h , deren Argumente Verteilungen auf A sind, heißt stetig, falls aus $\varphi_n \rightarrow 0$ auch $h\varphi_n \rightarrow 0$ folgt³⁾.

In obigen Bezeichnungen sei Φ eine Menge von beschränkten Verteilungen auf A , die alle beschränkten stetigen Verteilungen auf A und alle Summen und Produkte ihrer Elemente enthält, h eine auf Φ erklärte (reelle) Homomorphie, d. h. aus $\varphi_0 \in \varphi_1 + \varphi_2$ bzw. $\varphi_0 \in \varphi_1 \cdot \varphi_2$ ⁴⁾ folgt $h\varphi_0 = h\varphi_1 + h\varphi_2$ bzw. $h\varphi_0 = h\varphi_1 \cdot h\varphi_2$, $h\varphi$ immer reell.

Satz 1. *Folgende Bedingungen sind einander gleichwertig:*

1. h ist stetig,
2. ist g eine stetige Funktion einer reellen Veränderlichen und $\varphi_0, \varphi_1 \in \Phi$; $\varphi_0 \in g(\varphi_1)$, so ist auch $h\varphi_0 = g(h\varphi_1)$,
3. ist g eine stetige Funktion von N reellen Veränderlichen und $\varphi_0, \varphi_1, \dots, \varphi_N \in \Phi$; $\varphi_0 \in g(\varphi_1, \dots, \varphi_N)$, so ist $h\varphi_0 = g(h\varphi_1, \dots, h\varphi_N)$,
4. ist j eine offene Umgebung von $h\varphi$, $\varphi \in \Phi$, so ist $\varphi(j)$ von Null verschieden.

Die letzte Bedingung (mit der Homomorphieeigenschaft von h) besagt: h ist ein Charakter von Φ im Sinne von MF., S. 333; und h ist stetig im Sinne von MF., I. c.

In den Verabredungen von MF., S. 333 sei F die Menge der beschränkten meßbaren Funktionen, k eine auf F erklärte (reelle) Homomorphie: $k(f_1 + f_2) = kf_1 + kf_2$, $k(f_1 f_2) = kf_1 \cdot kf_2$. Zwischen zwei gleichen (d. h. nur auf einer Nullmenge voneinander abweichenden) Funktionen f_1 und f_2 wollen wir gar nicht unterscheiden; insbesondere $kf_1 = kf_2$. Ist nun $f_0 \in g(f_1, \dots, f_N)$ mit $f_0 = g(f_1, \dots, f_N)$ gleichbedeutend und erklärt man die Stetigkeit von k ebenso wie für h , wo $|f| \leq \varepsilon$ mit den Ungleichungen $|f(x)| \leq \varepsilon$ für alle x gleichbedeutend ist, so gilt

Satz 1a. *Satz 1 bleibt bestehen, wenn man Φ , φ und h durch F , f und k ersetzt.*

Wie in MF., S. 333 nenne ich jede Funktion k , die Satz 1a befriedigt, einen Charakter von F .

Jeder meßbaren Funktion f ist kraft MF., S. 333 eine Verteilung φ zugeordnet. Laut MF. II. Satz nebst Zusatz entsprechen sich h und k ein-

³⁾ Man muß dies vom Stetigkeitsbegriff für Charaktere, MF., S. 333, unterscheiden! Erst Satz 1 zeigt, daß es für Homomorphismen auf dasselbe hinauskommt.

⁴⁾ Siehe MF. II, S. 332.

deutig kraft der Gleichung $h\varphi = k\varphi$. Ist also \mathfrak{B} eine der Bedingungen von Satz 1 bzw. 1a, so ergibt sich

Folgesatz. Die Homomorphismen h von Φ , die \mathfrak{B} erfüllen, entsprechen den \mathfrak{B} befriedigenden Homomorphismen k von F eindeutig: $k\varphi = h\varphi$.

Es genügt den Beweis von Satz 1 durchzuführen. Der Beweis von Satz 1a verläuft ganz parallel; gewisse Einzelheiten ergeben sich dabei als ganz nahelegend.

Beweis von Satz 1.

g mit etwaigen Indizes sei immer eine stetige Funktion einer reellen Veränderlichen derart, daß die x -Achse aus endlichvielen Monotonieintervallen von $g(x)$ besteht, $\varphi_s(j) = \varphi(g^{-1}(j))$.

(a) Nach (α) und (β) ist φ_s eine Verteilung auf A . Ist φ stetig, bzw. beschränkt, so ist dies auch für φ_s der Fall. Ferner gilt $\varphi_s \in g(\varphi)$ und $|\varphi_s| \leq \sup g(x) (x \in j)$, wo $\varphi(j) = \varepsilon$ ist.

Ist ferner $g \in G(g', g'')$, G stetig, so gilt

$$(b) \quad \varphi_s \in G(\varphi_{s'}, \varphi_{s''}).$$

Denn ist $j = G(j_1, j_2)$, so ist $\varphi_{s'}(j_1) \varphi_{s''}(j_2) = \varphi(g'^{-1}(j_1)) \varphi(g''^{-1}(j_2))$, also nach (β) gleich $\varphi(g'^{-1}(j_1) \cdot g''^{-1}(j_2)) \subset \varphi(g^{-1}(j)) = \varphi_s(j)$.

Für jede reelle Zahl r setzen wir nun $\varphi_r(j)$ gleich ε bzw. gleich Null, je nachdem j die Zahl r enthält oder nicht. φ_r ist ja beschränkt und stetig und man hat

$$(c) \quad \varphi_g(r, s, \dots, z) \in g(\varphi_r, \varphi_s, \dots, \varphi_z).$$

Ferner gilt immer $\varphi \in \varphi_1 \varphi$, also $h\varphi = h\varphi_1 \cdot h\varphi$, woraus sich $h\varphi_1 = 1$ ergibt. (Wir schließen den trivialen Fall aus, wo h identisch verschwindet.) Ist nun $r = p/q$, p und q natürliche Zahlen, so ist $h\varphi_p = q \cdot h\varphi_r$. (Ist nämlich $q-1 > 0$, $p' = r(q-1)$ und setzt man $h\varphi_{p'} = (q-1) \cdot h\varphi_r$, als bewiesen voraus, so gilt nach (c) $\varphi_p \in \varphi_{p'} + \varphi_r$, also $h\varphi_p = h\varphi_{p'} + h\varphi_r = q \cdot h\varphi_r$.) Ebenso gilt $h\varphi_p = p \cdot h\varphi_1 = p$, also schließlich $h\varphi_r = r$.

Ferner gilt $\varphi_0 = \varphi_2 \varphi_0$, d. h. $h\varphi_0 = h\varphi_2 \cdot h\varphi_0 = 2 h\varphi_0$, also $h\varphi_0 = 0$. Nach (c) gilt schließlich $\varphi_0 \in \varphi_{-r} + \varphi_r$, also $0 = h\varphi_{-r} + h\varphi_r$, d. h. $h\varphi_{-r} = -r$. Also

$$(d) \quad h\varphi_r = r \text{ für jede rationale Zahl } r.$$

(e) Ist g ein Polynom mit rationalen Koeffizienten, so gilt immer $\varphi_g \in \Phi$, $h\varphi_g = g(h\varphi)$ ($\varphi \in \Phi$).

Denn ist erstens g identisch gleich Null, so ist offenbar $\varphi_g = \varphi_0$ und unsere Behauptung kraft (d) richtig. Unser $g(x)$ kann offenbar aus Null durch Additionen rationaler Konstanten und Multiplikationen mit x aufgebaut werden. Also sei $G(u, v) = u + v$, g' eine rationale Konstante und g'' ein Polynom mit $g''(0) = 0$ oder aber $G(u, v) = uv$, $g'(x) = x$ und g'' ein Polynom

niedrigeren Grades als g . Es gelte immer $g = G(g', g'')$. Für g'' darf man (e) als schon bewiesen annehmen. Aus $\varphi \in \Phi$ folgt ja $\varphi_{g'} \in \Phi$ und nach Induktionsvoraussetzung $\varphi_{g''} \in \Phi$. Aus (b) ergibt sich nun $\varphi_g \in G(\varphi_{g'}, \varphi_{g''})$, also kraft unserer Annahmen über Φ und h und des speziellen Charakters der Funktion G : $\varphi_g \in \Phi$ und $h\varphi_g = G(h\varphi_{g'}, h\varphi_{g''}) = G(g'(h\varphi), g''(h\varphi)) = g(h\varphi)$, w. z. b. w.; vgl. dazu (d) und (e).

(f) Ist h stetig, so gibt es zu jedem $\varepsilon > 0$ ein $\delta_\varepsilon > 0$, $\delta_\varepsilon \leq \varepsilon$ mit $|h\varphi| < \varepsilon$, falls nur $|\varphi| \leq \delta_\varepsilon$ gilt (und umgekehrt).

Wäre dies nicht der Fall, so gäbe es nämlich ein $\varepsilon > 0$ und eine Folge φ_n mit $|h\varphi_n| \leq \varepsilon$ und $|\varphi_n| \leq 1/n$, was der Stetigkeitsdefinition widerspricht.

(g) Ist h stetig und $\varphi, \varphi_g \in \Phi$, so gilt $h\varphi_g = g(h\varphi)$.

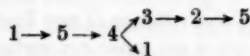
Zum Beweise sei $M > 0$ so gewählt, daß $|\varphi| \leq M$ und $|h\varphi| < M$ gilt; j sei das Intervall aller x mit $|x| \leq M$. Ist $\eta > 0$ beliebig, so kann man nach Weierstraß ein Polynom p mit $|g - p| < \eta$ auf j wählen, dessen Koeffizienten ja rational angenommen werden dürfen. Man setze noch $g' = g - p$. Man hat $\varphi_g \in \Phi$, $\varphi_{-p} \in \Phi$ kraft (e) und nach (b): $\varphi_{g'} \in \varphi_g + \varphi_{-p}$, also $\varphi_{g'} \in \Phi$ und $h\varphi_{g'} = h\varphi_g + h\varphi_{-p} = h\varphi_g - h\varphi_p$; denn $\varphi_p \in \Phi$ kraft (e) und $\varphi_p \in \varphi_{-1} \cdot \varphi_{-p}$ nach (b), also $h\varphi_p = -h\varphi_{-p}$ laut (d).

Nun sei ein beliebiges $\varepsilon = 2\varrho > 0$ gegeben, $\eta = \delta_\varrho$ in den Bezeichnungen von (f). Aus $|g'| < \eta$ und (a) erhält man $|\varphi_{g'}| \leq \eta$, d. h. $|h\varphi_g - h\varphi_p| = |h\varphi_{g'}| < \varrho$. Kraft (e) kann man dies $|h\varphi_g - p(h\varphi)| < \varrho$ schreiben. Außerdem ist ja $|g(h\varphi) - p(h\varphi)| < \eta \leq \varrho$, also $|g(h\varphi) - h\varphi_g| < 2\varrho = \varepsilon$ für jedes $\varepsilon > 0$, d. h. $g(h\varphi) = h\varphi_g$, w. z. b. w.

Neben den Bedingungen von Satz 1 betrachten wir noch folgende:

5. Ist $\varphi, \varphi_g \in \Phi$, so gilt $h\varphi_g = g(h\varphi)$.

Im Beweise der Gleichwertigkeit der Bedingungen 1 bis 5 will ich nach folgendem Schema verfahren:



Dabei gilt $1 \rightarrow 5$ nach (g), $4 \rightarrow 3$ nach MF. II, Satz, $3 \rightarrow 2$ ist trivial und $2 \rightarrow 5$ gilt nach (a). Auch $4 \rightarrow 1$ liegt nahe. j_1 sei nämlich das abgeschlossene Intervall $< -\varepsilon, \varepsilon >$, j_1' gleich $j_2 + j_3$, $j_2 j_3$ leer, $\varphi(j_1 + j_2) = \varepsilon + \varphi(j_2)$, $\varphi(j_2) = \varphi((j_1 + j_2)j_2) = (\varepsilon + \varphi(j_2)) \cdot \varphi(j_2) = 0$. Gilt also 4, so ist $h\varphi$ non $\in j_2$ und ebenso $h\varphi$ non $\in j_3$, d. h.

(h) aus $|\varphi| \leq \varepsilon$ folgt $|h\varphi| \leq \varepsilon$,

woraus die zu beweisende Stetigkeit von h folgt.

Es bleibt also übrig, $5 \rightarrow 4$ zu beweisen. Wir setzen also 5 als gültig voraus. Angenommen, 4 wäre falsch, so gibt es ein offenes j_0 und eine Verteilung $\varphi \in \Phi$ mit $h\varphi \in j_0$, $\varphi(j_0) = 0$. Die Endpunkte r und s von j_0 dürfen rational angenommen werden, $r < h\varphi < s$. Ist nun $x < r$ bzw. $x > s$, so

setze ich $g(x) = r$ bzw. $g(x) = s$ und $g(x) = x$ für $r \leq x \leq s$. Offenbar ist g stetig und φ_g hat folgende charakteristische Eigenschaft ($\varrho = \varphi_g(r)$, $\sigma = \varphi_g(s)$): enthält j weder r noch s , so ist $\varphi_g(j) = 0$, $\varrho + \sigma = e$. Nach 5 ist ferner $h\varphi_g = h\varphi \neq r, s$. Man hat also nur zu zeigen, daß $h\varphi_g$ entweder r oder s gleich ist. Man setze $\lambda = \varphi_g(l)$, l reell. Man setze ferner $\varphi'(1) = \lambda$, $\varphi'(0) = \lambda'$ und $\varphi'(j) = 0$, falls j weder Null noch Eins enthält. Dann gilt ja $\varphi' \in \varphi' \cdot \varphi'$, d. h. $h\varphi' = 0$ oder 1. Ferner sieht man leicht $\varphi_1 \in \varphi' + \varphi'$, also $h\varphi' + h\varphi' = 1$. Es sei nun u diejenige der Zahlen r und s , für die $h\varphi^u = 1$ ist, und v die andere, also $h\varphi^v = 0$. Es genügt zu zeigen, daß $h\varphi = u$ ist.

Nun sei ${}_1\varphi(l) = \lambda$, ${}_1\varphi(0) = \lambda'$ und ${}_1\varphi(j) = 0$ für $0, l \text{ non } \in j$. Dann ergibt sich leicht ${}_1\varphi \in \varphi_1 \cdot \varphi'$, also $h{}_1\varphi = h\varphi_1 \cdot h\varphi' = l \cdot h\varphi'$ nach (d), falls l rational ist. Ferner sieht man ohne Mühe, daß $\varphi_g \in r\varphi + s\varphi$, also $h\varphi_g = h_r\varphi + h_s\varphi = h_r\varphi + h_s\varphi = u h\varphi^u + v h\varphi^v = u$, w. z. b. w.

Damit ist Satz 1 völlig bewiesen.

II.

Die Bairesche Eigenschaft.

Die Norm $\|\varphi\|$ von φ sei die untere Grenze aller $M > 0$ mit $|\varphi| \leq M$. Sind φ_1, φ_2 stetige Verteilungen, so darf man $\varphi = \varphi_1 + \varphi_2$ und $\varphi = \varphi_1 \cdot \varphi_2$ statt $\varphi \in \varphi_1 + \varphi_2$ und $\varphi \in \varphi_1 \cdot \varphi_2$ schreiben, da nach MF. II, Bemerkung 2 die Summen und Produkte im Bereiche der stetigen Verteilungen eindeutig bestimmt sind.

Φ sei nun die Gesamtheit aller beschränkten stetigen Verteilungen auf A . Wir fragen, unter welchen Bedingungen es eine Isomorphie von Φ mit einer Teilmenge des Ringes C der beschränkten Zahlenfolgen gibt, die die Norm unverändert läßt⁵⁾. Wir wollen dann $\|\Phi\| \subset \|C\|$ schreiben. Dazu wird folgende Begriffsbildung von Nutzen sein. Unser Boolescher Ring A heißt *separabel*, falls es eine abzählbare Menge Π von Primidealen in A gibt derart, daß für jedes $a \in A$ in Π ein π mit $a \text{ non } \in \pi$ vorhanden ist (a von Null verschieden angenommen).

Im Beweise von Satz 2 wird eine von Stone herrührende topologische Darstellung von A von Nutzen sein⁶⁾. Unter einem *Baustein* verstehe ich eine zugleich offene und abgeschlossene Menge. Nach Stone (l. c.) gibt es einen bikompakten Hausdorffschen Raum $e = \mathfrak{E}(A)$, in dem jede offene Menge eine Vereinigung von Bausteinen ist, derart, daß A (von Isomorphismen abgesehen) der Mengenkörper aller Bausteine in e ist. e ist die Einheit von A . Für Punkte π von e kann man die Primideale π in A nehmen, und zwar derart, daß $\pi \in a$ mit $a \text{ non } \in \pi$ gleichbedeutend ist.

⁵⁾ $\|\{a_n\}\| =$ obere Grenze aller a_n .

⁶⁾ M. H. Stone, Trans. Amer. Math. Soc. 41 (1937), Theorem 1 und 2.

Nach MF. II, Satz nebst Zusatz gibt es ferner eine eindeutige Zuordnung zwischen den Charakteren h von Φ und den Primidealen π in A , $h = h_\pi$, derart, daß aus $h\varphi \in j$, j offen, folgt: $\varphi(j) \text{ non } \in \pi$.

Eine Verteilung φ heißt *elementar*, falls es eine endliche Zahlenmenge Z derart gibt, daß $\varphi(j)$ für jedes zu Z fremde j verschwindet.

Man setze noch $f(\pi) = h_\pi \varphi$.

(i) f ist auf e stetig, falls φ eine beschränkte stetige Verteilung auf A ist.

Ist φ elementar und Z wie oben erklärt, so ist offenbar $h_\pi \varphi = z$ mit $\varphi(z) \text{ non } \in \pi$ gleichwertig. Denn ist j offen, $z \in j$, $j \cdot (Z + z)$ leer, so ist ja $\varphi(j) \text{ non } \in \pi$, $\varphi(j) \varphi(Z + z) = 0 \in \pi$, also, da π prim ist, $\varphi(Z + z) \in \pi$. Wäre nun $\varphi(z) \in \pi$, so wäre $e = \varphi(Z + z) + \varphi(z) \in \pi$, was nicht der Fall ist. Also ist $f(\pi) = z$ mit $\pi \in \varphi(z)$ gleichbedeutend, also f stetig. Nach MF. VI, Satz 5 können nun die Funktionen f für beliebige φ mittels solcher für elementare Verteilungen gleichmäßig approximiert werden, woraus sich ihre Stetigkeit ergibt.

Satz 2. Es gilt dann und nur dann $\|\Phi\| \subset \|C\|$, wenn A separabel ist.

Beweis. Aus MF. VI, Satz 1 nebst Zusatz 2 und (h) folgt

$$(j) \quad \|\varphi\| = \|f\| \quad (f(\pi) = h_\pi \varphi), \quad \|f\| = \sup |f|^7).$$

$\Pi = \{\pi_n\}$ sei eine abzählbare, in e dichte Punktfolge. Für partielle Funktionen f' von f auf Π gilt laut (j): $\|\varphi\| = \|f'\|$ und die Zuordnung $\varphi \rightarrow f'$ ist ja homomorph.

(k) Setzt man $a_n = f(\pi_n)$, so ist $\varphi \rightarrow \{a_n\}$ eine die Inklusion $\|\Phi\| \subset \|C\|$ realisierende Zuordnung.

Denn die Korrespondenz $\varphi \rightarrow f$, also auch $\varphi \rightarrow f'$ ist nach MF. I, Hilfssatz 2 eineindeutig.

Umgekehrt gebe es eine solche Zuordnung $\varphi \rightarrow \{a_n\}$. Für jedes feste n ist dann $\varphi \rightarrow a_n$ eine stetige Homomorphie, also nach Satz 1 ein Charakter. Laut MF. II, Satz gibt es also ein Primideal $\pi = \pi_n$ in A mit $h_\pi \varphi = a_n$. Man hat zu zeigen, daß die Punkte π_n in e dicht liegen. Wäre dies nicht der Fall, so gäbe es einen Baustein $0 \neq a \in A$ mit $\pi_n \text{ non } \in a$ für alle n . Ferner sei $|\varphi| \leq M$, $z > M$. Ich setze $\varphi'(j) = a' \varphi(j)$, falls $z \text{ non } \in j$ und $\varphi'(j) = a$. Offenbar ist $\varphi' \in \Phi$ und $\varphi' \neq \varphi$. Es genügt zu zeigen, daß $h_\pi \varphi' = a_n$ für $\pi = \pi_n$ gilt. Oder stärker: Ist $a \in \pi$, so ist $h_\pi \varphi' = h_\pi \varphi$.

j sei also offen, $h_\pi \varphi \in j$. Man hat nur zu zeigen, daß $\varphi'(j)$ in π nicht enthalten ist. Nach (h) ist $h_\pi \varphi \neq z$, also kann man j so klein wählen, daß $z \text{ non } \in j$ ist. (Für größere j gilt es desto mehr.) Dann ist $\varphi'(j) = a' \varphi(j)$ mit $a' \text{ non } \in \pi$ und $\varphi(j) \text{ non } \in \pi$, also, da π prim ist, $\varphi'(j) \text{ non } \in \pi$, w. z. b. w.

⁷⁾ $\sup \varphi$ = obere Grenze aller Werte der Funktion φ .

Von jetzt an sei A immer der nach dem Ideal β aller Mengen der ersten Kategorie reduzierte Ring (Mengenkörper) B aller Mengen mit der Baireschen Eigenschaft³⁾ in einem Euklidischen Raume: $A = B/\beta = lB$, $l\beta = 0$.

(l) A ist separabel.

Zum Beweis möge o_n eine abzählbare offene Basis unseres Grundraumes durchlaufen. Den Mengen o_n mögen in A Elemente $a_n = lo_n$ entsprechen, die ja von Null verschieden sind. Es gibt also ein Primideal π_n in A , das a_n nicht teilt (vgl. St. 63). Ist nun a irgendein Element von A , $a \neq 0$, so gibt es eine offene Menge o , der a entspricht: $a = lo$. Ferner gibt es ein n mit $o_n \subset o$, woraus $a_n \subset a$, also auch $a \text{ non} \in \pi_n$ folgt, so daß $\Pi = \{\pi_n\}$ der Separabilitätsdefinition für A genügt.

Satz 3. Man hat $\|\Phi\| \subset \|C\|$.

Beweis. Dies folgt aus Satz 2 und (l).

F sei die Gesamtheit aller beschränkten, in B meßbaren Funktionen, wobei man unter Funktionen, die nur auf einer Menge aus β voneinander abweichen, gar nicht unterscheidet. Mit $\|f\|$, $f \in F$ bezeichnen wir die untere Grenze aller M , für die $|f| \leq M$ bis auf eine Menge aus β ist. Jeder $f \in F$ ist eine Verteilung $\varphi(j) = lf^{-1}(j)$ wie in MF., S. 333 zugeordnet.

(m) Die Zuordnung $f \rightarrow \varphi$ ist eine Isomorphie, die die Norm nicht ändert.

Die Isomorphieeigenschaft folgt aus MF. II, Hilfssatz 1 und IX, Satz und die Gleichheit der Normen liegt nahe.

Zusatz. Man hat $\|F\| \subset \|C\|$.

Beweis. Dies folgt aus Satz 3 und (m).

III.

Zahlenfolgen.

\mathfrak{E} sei ein Mengenkörper, F die Vereinigung aller Mengen aus \mathfrak{E} , $E \in \mathfrak{E}$, e ein Ideal in \mathfrak{E} , deren Elemente für Nullmengen genommen werden mögen, $A = \mathfrak{E}/e = l\mathfrak{E}$, $e = lE$. Die beschränkten, in \mathfrak{E} meßbaren Funktionen mögen einen Ring R bilden. Für $r \in R$ schreibe ich $r \equiv 0$, falls aus $0 \text{ non} \in j, j$ abgeschlossen, die Inklusion $r^{-1}(j) \in e$ folgt. $r_1 \equiv r_2$ besagt: $r_1 - r_2 \equiv 0$.

(n) Dann und nur dann ist $r_1 \equiv r_2$, wenn alle Charaktere k von R für r_1 und r_2 miteinander übereinstimmen: $kr_1 = kr_2$.

Setzt man nämlich $r = r_1 - r_2$, so folgt aus $r \equiv 0$ sicher $kr = 0$, also $kr_1 = kr_2$.

³⁾ $b \in B$ besagt: $b = o + z$, o offen, $z \in \beta$. Φ ist ein Ring kraft MF. V, Satz 3 oder kraft (m). Aus (j) und MF. V, Satz 4 folgt weiter, daß Φ und nach (m) auch F ein Banachscher Raum ist.

Ist umgekehrt $r \neq 0$, so gibt es ein abgeschlossenes j mit $0 \text{ non } \in j$, $r^{-1}(j) \text{ non } \in e$, d. h. $\eta = lr^{-1}(j) \neq 0$. Nun sei $e_0(j^*) = 0$ bzw. $= \varepsilon$, je nachdem $0 \text{ non } \in j^*$ oder $0 \in j^*$ gilt. j^* sei offen, $0 \in j^*$, jj^* leer. Dann gilt ja $e_0(j^*) = \varepsilon$, $e(j^*) \subset \varepsilon + \eta$ ($e = lr^{-1}$), also $e_0(j^*) \text{ non } \subset e(j^*)$. Nach MF. I, Hilfssatz I und II, Satz stimmen die Charaktere von e mit denen von e_0 nicht überein. Da nun für e_0 jeder Charakter verschwindet, so gibt es ein h mit $h e \neq 0$, also laut MF. II, Satz ein k mit $kr \neq 0$, w. z. b. w.

(o) \equiv ist eine Kongruenz (d. h. reflexiv, symmetrisch und transitiv) und die Kongruenzen darf man addieren und multiplizieren.

Dies folgte aus (n).

(R) sei die Gesamtheit aller Klassen (r) , wo $r' \in (r)$ mit $r' \equiv r$ gleichbedeutend ist. In naheliegender Weise kann (R) laut (o) als ein Ring aufgefaßt werden. Mit $|(r)|$ bezeichnen wir die untere Grenze aller $\supr |r'|$ mit $r' \equiv r$. k möge immer alle Charaktere von R durchlaufen.

(p) $|(r)|$ ist die obere Grenze aller kr .

Erstens sei $|(r)| < M$. Dann gibt ein $r' \equiv r$ mit $\supr |r'| < M$, also $|kr'| \leq M$, d. h. laut (n): $|kr| \leq M$, also $|kr| \leq |(r)|$ für alle k .

Umgekehrt sei $|kr| < M$ für alle k , j abgeschlossen und $|x| \geq M$ für alle $x \in j$. Ist $\eta = lr^{-1}(j)$ von Null verschieden, so gibt es nach St. 63 ein Primideal π in A , $\eta \text{ non } \in \pi$. Laut MF. II, Satz nebst Zusatz gehört zu π ein k . Kraft Voraussetzung hat man $kr \text{ non } \in j$. Es gibt also ein offenes j^* , $kr \in j^*$, jj^* leer. Ferner ist $e(j^*) \text{ non } \in \pi$ ($e = lr^{-1}$) und $e(j) = \eta \text{ non } \in \pi$, also $0 = e(jj^*) = e(j)e(j^*) \text{ non } \in \pi$, was ja nicht geht.

Also ist $r^{-1}(j) \in e$ und man kann leicht auf der Nullmenge $E|_x |r(x)| \geq M$ die Funktion derart modifizieren, daß die entstehende Funktion r' der Ungleichung $\supr |r'| \leq M$ genügt. Dann ist $r' \equiv r$, also $|(r)| \leq M$, w. z. b. w.

Nimmt man statt $|(r)|$ die untere Grenze $\|r\|$ aller M mit $|r| \leq M$ bis auf eine Nullmenge und für \equiv die Gleichheit bis auf Nullmengen, so bleibt die obige Überlegung bestehen, also

(p') $\|r\|$ ist die obere Grenze aller kr , also nach (p) gleich $|(r)|$.

Von jetzt an schreibe ich also $\|r\|$ statt $|(r)|$. Der Funktion $r \in R$ entspricht nun eine Funktion f auf $\varepsilon = \mathfrak{E}(A)$ (vgl. Vorrede zu Satz 2): $f(\pi) = kr$, wo k laut MF. II, Satz nebst Zusatz zu π gehört.

(i') f ist auf ε stetig.

Denn jede Funktion $r \in R$ kann durch elementare (d. h. nur endlich viele Werte annehmende) Funktionen aus R gleichmäßig approximiert werden, woraus sich nach (p) eine ähnliche Approximation für die zugehörige Funktion f ergibt. Und für elementare r ergibt sich (i') aus (i) und der Korrespondenz $k \rightarrow h_\pi$ (vgl. MF., I c.).

Aus (p) folgt weiter

$$(j') \quad ||r|| = ||f|| = \sup |f|.$$

Satz 2a. *Es gilt dann und nur dann $||(R)|| \subset ||C||$, wenn A separabel ist.*

Beweis. Liegt die Folge $\Pi = \{\pi_n\}$ dicht in ε , so gilt laut (j'): $||r|| = ||f'||$, wo f' die auf Π erklärte partielle Funktion von f ist. Ferner ist die Zuordnung $(r) \rightarrow f'$ homomorph und eineindeutig. Denn ist f' , also nach (i') auch f identisch Null, so ist laut (n) und (j') auch $(r) = 0$. Also

(k') setzt man $a_n = f(\pi_n)$, so ist $(r) \rightarrow \{a_n\}$ eine die Inklusion $||(R)|| \subset ||C||$ realisierende Zuordnung.

Umgekehrt sei $(r) \rightarrow \{a_n\}$ eine solche Zuordnung. Dann ist $r \rightarrow a_n$ (n fest) laut Satz 1a ein Charakter $k: kr = a_n$; $\pi = \pi_n$ sei das laut MF. II, Satz nebst Zusatz zu k gehörige Primideal in A . Würden nun die Punkte π_n in ε nicht dicht liegen, so gäbe es eine Menge $o \in \mathfrak{E}$, o non $\in e$, $o = lo \in \pi_n$ für alle n . r' sei nun überall gleich r , nur auf o sei r' gleich $M > \sup |r|$, $r' \in R$. Dann ist nach (p') $(r') \neq (r)$; denn $||r'|| = M$.

Um den gesuchten Widerspruch zu gewinnen, genügt es also, $kr' = kr$ für $\pi = \pi_n$ (vgl. MF., l. c.) zu beweisen, also $\eta = lr'^{-1}(j)$ non $\in \pi$, j offen, $kr \in j$. Man darf annehmen: M non $\in j$. Dann ist $\eta = o' \cdot lr^{-1}(j)$. Da nun keiner der Faktoren durch das Primideal π teilbar ist, so gilt dies auch von η , w. z. b. w.

Jetzt sei \mathfrak{E} der Ring aller Mengen von ganzen Zahlen, e das Ideal aller endlichen Mengen darin. Die meßbaren Funktionen sind also die beschränkten Zahlenfolgen und $||r|| = 0$ besagt: r ist eine Nullfolge.

Satz 3a. *Man hat $||(R)||$ non $\subset ||C||$.*

Beweis. Nach Satz 2a hat man zu zeigen:

(l') A ist nicht separabel⁹⁾.

Zum Beweise darf man annehmen, \mathfrak{E} sei der Ring aller Mengen rationaler Zahlen. Zu jeder Irrationalzahl i wählen wir eine rationale Folge (i) , die i zustrebt. Dann ist immer $l(i)$ von Null verschieden und die $l(i)$ annullieren sich zu je zweien. Also gibt es in ε un abzählbar viele zu je zweien fremde Bausteine $l(i)$, so daß in ε keine abzählbare Folge dicht liegt.

P sei die Gesamtheit aller beschränkten stetigen Verteilungen auf A .

Zusatz. *Man hat $||P||$ non $\subset ||C||$.*

Beweis. Dies folgt aus Satz 2 nebst (l').

⁹⁾ Vgl. Theorem IV meiner Abhandlung „On bicompact spaces“, Publications de la Faculté des Sciences de l'Université Masaryk à Brno 270 (1939), S. 8.

Über Ketten von Faktoroiden.

Von

O. Borůvka in Brünn.

In den Vorlesungen über Gruppentheorie, die ich im Jahre 1938/39 auf der Universität in Brünn gehalten habe, war ich bestrebt, die allgemeinen Sätze der Gruppentheorie mit Hinblick auf ihre Abhängigkeit von den einzelnen Gruppenaxiomen zu formulieren und zu beweisen. Das hat mich naturgemäß zu systematischen Untersuchungen über Gruppoide geführt. Unter einem Gruppoid verstehe ich den Inbegriff einer nicht leeren Menge G und einer Multiplikation in G . In dieser Arbeit erlaube ich mir einen Teil meiner Resultate darzustellen, und zwar die Theorie der Ketten von Faktoroiden. Der Begriff eines Faktoroides ist für die gesamte Gruppoidentheorie von grundlegender Bedeutung und stellt eine Verallgemeinerung des Begriffes einer Faktorgruppe dar. Ich erlaube mir auf seine Definition in Nr. 8 hinzuweisen. Die hier entwickelte Theorie der Ketten von Faktoroiden bildet eine weitgehende Verallgemeinerung der klassischen Theorie der Normalketten.

Die Arbeit besteht aus drei Teilen. Im Teile I werden mengentheoretische Grundlagen der Gruppoidentheorie entwickelt; der Teil II enthält grundlegende Begriffe und Sätze der Gruppoidentheorie und der Teil III die Theorie der Ketten von Faktoroiden.

I. Mengentheoretische Grundlagen der Gruppoidentheorie.

1. **Bezeichnungen.** Mengen (Elemente in Mengen) bezeichnen wir in der Regel mit großen (kleinen) lateinischen Buchstaben. Systeme von Mengen bezeichnen wir in der Regel mit Symbolen wie z. B. \bar{A} , \bar{A} und ihre einzelnen Elemente mit \bar{a} , \bar{a} . $s\bar{A}$ bedeutet die Summe aller Mengen, die in \bar{A} als Elemente vorkommen. 0 ist das Symbol für die leere Menge. Die auf zwei Mengen sich beziehenden Symbole \vee , \wedge , \cap bedeuten ihre Summe, Differenz, Durchschnitt. Die Bedeutung von \in , \subset , \supset , $=$ und \notin , $\not\subset$, $\not\supset$ dürfen wir wohl als bekannt voraussetzen. Wenn zwei Mengen einen nicht leeren Durchschnitt haben, so nennen wir sie *inzident*. Der Buchstabe G bedeutet die ganze Arbeit hindurch eine nicht leere Menge.

2. **Zerlegungen in Mengen.** Ein nicht leeres System \bar{A} von nicht leeren paarweise fremden Untermengen in G nennen wir *Zerlegung in G* und im Falle $s\bar{A} = G$ *Zerlegung von G oder auf G* . Ist \bar{A} eine Zerlegung in G und \bar{A} eine Zerlegung von \bar{A} , so ist das System \bar{A} von Untermengen in G , von denen jede die Summe aller in demselben Elemente von \bar{A} vorkommenden Elemente von \bar{A} ist, wieder eine Zerlegung in G . Wir sagen, \bar{A} sei die durch \bar{A} erzeugene Überdeckung von \bar{A} , kürzer: eine *Überdeckung von \bar{A}* .

Es seien \bar{A}, \bar{C} Zerlegungen in G . Die Menge aller nicht leeren Durchschnitte eines jeden Elementes in \bar{A} mit jedem Elemente in \bar{C} heißt die *Durchdringung von \bar{A}, \bar{C}* und wird mit $\bar{A} \cap \bar{C}$ oder $\bar{C} \cap \bar{A}$ bezeichnet. Es ist $s(\bar{A} \cap \bar{C}) = s\bar{A} \cap s\bar{C}$. In ähnlicher Weise definieren wir die *Durchdringung einer nicht leeren Untermenge $B \subset G$ und der Zerlegung \bar{A}* , und zwar als die Menge aller nicht leeren Durchschnitte von B mit den Elementen in \bar{A} . Bezeichnung: $B \cap \bar{A}$ oder $\bar{A} \cap B$.

Die Menge der Elemente $\bar{c} \in \bar{C}$, die mit irgendeinem Elemente $\bar{a} \in \bar{A}$ inzident sind, nennen wir die *Hülle der Zerlegung \bar{A} in \bar{C}* und bezeichnen sie mit $\bar{A} \subset \bar{C}$ oder $\bar{C} \supset \bar{A}$. Nach dieser Definition ist also $\bar{A} \subset \bar{C} \subset \bar{C}$ und es gilt $s\bar{A} \cap s(\bar{A} \subset \bar{C}) = s\bar{A} \cap s\bar{C}$. In ähnlicher Weise definieren wir die *Hülle einer nicht leeren Untermenge $B \subset G$ in der Zerlegung \bar{A}* , und zwar als die Menge aller Elemente $\bar{a} \in \bar{A}$, die mit B inzident sind. Bezeichnung: $B \subset \bar{A}$ oder $\bar{A} \supset B$.

Es ist zweckmäßig, die Symbole $\bar{A} \cap \bar{C}$, $B \cap \bar{A}$, $\bar{A} \subset \bar{C}$, $B \subset \bar{A}$ auch für den Fall zu definieren, wo ein oder beide Buchstaben in einem solchen Symbole die leere Menge 0 bedeuten, und zwar wieder als 0.

Der Begriff der Durchdringung von zwei Zerlegungen hängt mit dem Begriffe der Durchdringung einer Untermenge und einer Zerlegung und ebenso der Begriff der Hülle einer Zerlegung in einer weiteren Zerlegung mit dem Begriffe der Hülle einer Untermenge in einer Zerlegung in folgender Weise zusammen:

$$\bar{A} \cap \bar{C} = (s\bar{A} \cap \bar{C}) \cap (\bar{A} \cap s\bar{C}), \quad \bar{A} \subset \bar{C} = s\bar{A} \subset \bar{C}.$$

Wenn eine Untermenge B und eine Zerlegung \bar{A} in G gegeben sind, wobei $B \cap s\bar{A} \neq 0$, so ist dadurch einerseits die Zerlegung $B \cap \bar{A}$ von $B \cap s\bar{A}$ und andererseits die nicht leere Untermenge $B \subset \bar{A}$ in \bar{A} eindeutig bestimmt.

Wir überlassen es dem Leser, die folgenden einfachen Beziehungen zwischen Durchdringungen und Hüllen zu beweisen. Dabei bedeuten \bar{A}, \bar{C} Zerlegungen und $B \supset D$ Untermengen in G .

1. $1^\circ B \cap \bar{A} = (B \cap s\bar{A}) \cap \bar{A}$ $2^\circ B \subset \bar{A} = (B \cap s\bar{A}) \subset \bar{A}$;
2. $s(s\bar{A} \subset \bar{C}) \cap \bar{A} = s\bar{C} \cap \bar{A}$;
3. $(D \subset \bar{A}) \cap B = D \subset (\bar{A} \cap B) (\equiv D \subset \bar{A} \cap B)$. Daraus folgt (für $D = B$)
4. $(B \subset \bar{A}) \cap B = \bar{A} \cap B$;
5. $(B \subset \bar{A}) \cap D = \bar{A} \cap D$.

3. Zerlegungen auf Mengen. Wir werden nun insbesondere Zerlegungen auf G betrachten. Als Beispiele von solchen Zerlegungen führen wir an die aus dem einzigen Elemente $\{G\}$ bestehende *größte Zerlegung* \bar{G}_{\max} von G und die *kleinste Zerlegung* \bar{G}_{\min} von G ; diese besteht aus den Untermengen $\{a\}$, wobei $a \in G$.

Es seien \bar{G}_1, \bar{G}_2 Zerlegungen von G . Wenn jedes Element in \bar{G}_1 die Summe von einigen in \bar{G}_2 vorkommenden Elementen ist, so nennen wir $\bar{G}_1(\bar{G}_2)$ *Überdeckung (Verfeinerung) von $\bar{G}_2(\bar{G}_1)$* und sagen, $\bar{G}_1(\bar{G}_2)$ *sei oder liege über (unter) $\bar{G}_2(\bar{G}_1)$* . Bezeichnung: $\bar{G}_1 \geq \bar{G}_2$ oder $\bar{G}_2 \leq \bar{G}_1$.

1. Wenn $\bar{G}_1 \geq \bar{G}_2$ gilt, so folgt aus $\bar{a}_1 \in \bar{G}_1$, $\bar{a}_2 \in \bar{G}_2$, $\bar{a}_1 \cap \bar{a}_2 \neq 0$ die Beziehung $\bar{a}_1 \supset \bar{a}_2$ und umgekehrt.

Beweis. Ist $\bar{G}_1 \geq \bar{G}_2$ und $\bar{a}_1 \in \bar{G}_1$, $\bar{a}_2 \in \bar{G}_2$, $\bar{a}_1 \cap \bar{a}_2 \neq 0$, so ist \bar{a}_1 die Summe von einigen in \bar{G}_2 vorkommenden Elementen; unter diesen befindet sich das Element \bar{a}_2 wegen $\bar{a}_1 \cap \bar{a}_2 \neq 0$ und weil die Elemente von \bar{G}_2 paarweise fremd sind. Folgt umgekehrt aus $\bar{a}_1 \in \bar{G}_1$, $\bar{a}_2 \in \bar{G}_2$, $\bar{a}_1 \cap \bar{a}_2 \neq 0$ die Beziehung $\bar{a}_1 \supset \bar{a}_2$, so ist das Element \bar{a}_1 die Summe aller mit ihm inzidenten Elemente in \bar{G}_2 .

Wir bemerken, daß die oben definierte Beziehung \geq zwischen Zerlegungen von G reflexiv, antisymmetrisch und transitiv ist, also in jedem nicht leeren Systeme von Zerlegungen von G eine partielle Ordnung bestimmt. Wegen $\bar{G}_{\max} \geq \bar{G}_1 \geq \bar{G}_{\min}$ hat das durch diese Beziehung \geq teilweise geordnete System aller Zerlegungen von G das größte Element \bar{G}_{\max} und das kleinste Element \bar{G}_{\min} .

Wir betrachten nun ein nicht leeres System (\bar{X}) von Zerlegungen von G . Wenn für eine Zerlegung \bar{G} von G die Beziehung $\bar{G} \geq \bar{X}$ ($\bar{G} \leq \bar{X}$) besteht, und zwar für jede Zerlegung $\bar{X} \in (\bar{X})$, so nennen wir \bar{G} *Überdeckung (Verfeinerung) von (\bar{X})* oder *gemeinsame Überdeckung (Verfeinerung) von Zerlegungen in (\bar{X})* und sagen, \bar{G} *sei oder liege über (unter) (\bar{X})* . Eine Überdeckung von (\bar{X}) heißt *größte (kleinste) Überdeckung von (\bar{X})* , wenn sie über (unter) jeder Überdeckung von (\bar{X}) liegt; eine Verfeinerung von (\bar{X}) heißt *größte (kleinste) Verfeinerung von (\bar{X})* , wenn sie über (unter) jeder Verfeinerung von (\bar{X}) liegt. Offenbar ist \bar{G}_{\max} die einzige größte Überdeckung und \bar{G}_{\min} die einzige kleinste Verfeinerung von (\bar{X}) .

2. Es gibt genau eine kleinste Überdeckung von (\bar{X}) . Dieselbe ist durch die im Teile a) des folgenden Beweises gegebene Konstruktion bestimmt.

Beweis. a) Es sei $\bar{G}_0 \in (\bar{X})$ und $\bar{a}_0, \bar{b}_0 \in \bar{G}_0$. Eine geordnete endliche Menge von Elementen in \bar{G}_0

$$\{\bar{a}_1, \dots, \bar{a}_n\}$$

nennen wir *Kette* in (\bar{X}) von \bar{a}_0 nach \bar{b}_0 , wenn $\bar{a}_1 = \bar{a}_0$, $\bar{a}_n = \bar{b}_0$ und wenn je zwei einander folgende Elemente in der Menge mit einem und demselben Elemente einer geeigneten Zerlegung $\bar{X} \in (\bar{X})$ inzident sind. Die für je zwei Elemente \bar{a}_0, \bar{b}_0 in \bar{G}_0 durch die Existenz einer Kette in (\bar{X}) von \bar{a}_0 nach \bar{b}_0 definierte Beziehung ist offenbar reflexiv, asymmetrisch und transitiv. Es gibt also eine Zerlegung \bar{G}_0 von \bar{G}_0 , so daß für je zwei in demselben Element in \bar{G}_0 liegenden Elemente von \bar{G}_0 eine Kette in (\bar{X}) von dem einen nach dem anderen existiert, während es keine solche Kette gibt für zwei Elemente in \bar{G}_0 , die in verschiedenen Elementen von \bar{G}_0 liegen. Die durch \bar{G}_0 erzwungene Überdeckung \bar{U} von \bar{G}_0 ist die kleinste Überdeckung von (\bar{X}) .

b) \bar{U} ist eine Überdeckung von (X) . Zum Beweise betrachten wir eine Zerlegung $\bar{X} \in (\bar{X})$ von G und zwei inzidente Elemente $\bar{u} \in \bar{U}$, $\bar{x} \in \bar{X}$. Nach 1 genügt es zu zeigen, daß $\bar{u} \supset \bar{x}$ ist. Nun gibt es aber nach der Definition von \bar{u} und wegen $\bar{u} \cap \bar{x} \neq 0$ ein $\bar{a}_0 \in \bar{G}_0$, so daß $\bar{u} \supset \bar{a}_0$, $\bar{a}_0 \cap \bar{x} \neq 0$. Wir wählen ein $x \in \bar{x}$ und betrachten das durch die Beziehung $x \in \bar{b}_0$ bestimmte Element $\bar{b}_0 \in \bar{G}_0$. Offenbar ist $\{\bar{a}_0, \bar{b}_0\}$ eine Kette in (\bar{X}) von \bar{a}_0 nach \bar{b}_0 . Die Elemente \bar{a}_0, \bar{b}_0 von \bar{G}_0 sind also in demselben Elemente von \bar{G}_0 enthalten; also ist $\bar{u} \supset \bar{b}_0$. Daraus folgt $x \in \bar{u}$ und weiter $\bar{u} \supset \bar{x}$.

c) \bar{U} ist eine kleinste Überdeckung von (\bar{X}) . Zum Beweise betrachten wir eine Überdeckung \bar{G} von (\bar{X}) und zwei inzidente Elemente $\bar{a} \in \bar{G}$, $\bar{u} \in \bar{U}$. Nach 1 haben wir zu zeigen, daß $\bar{a} \supset \bar{u}$ ist. Nun ist aber \bar{a} die Summe von einigen Elementen in \bar{G}_0 und desgleichen \bar{u} . Daraus folgt wegen $\bar{a} \cap \bar{u} \neq 0$ die Existenz eines Elementes $\bar{a}_0 \in \bar{G}_0$, wobei $\bar{a}_0 \subset \bar{a} \cap \bar{u}$. Es sei $\bar{b}_0 \in \bar{G}_0$, $\bar{b}_0 \subset \bar{u}$. Es gibt eine Kette in (\bar{X}) von \bar{a}_0 nach \bar{b}_0

$$\{\bar{a}_1, \dots, \bar{a}_n\},$$

$$(\bar{a}_1 = \bar{a}_0, \bar{a}_n = \bar{b}_0)$$

und nach der Voraussetzung besteht die Beziehung $\bar{a}_1 \subset \bar{a}$. Wir nehmen an, daß für ein $(1 \leq) \beta (\leq \alpha - 1)$ auch die weiteren Beziehungen $\bar{a}_1, \dots, \bar{a}_\beta \subset \bar{a}$ bestehen. Dann gibt es in einer geeigneten Zerlegung $\bar{X} \in (\bar{X})$ von G ein mit $\bar{a}_\beta, \bar{a}_{\beta+1}$ inzidentes Element \bar{x} . Wegen $\bar{a}_\beta \subset \bar{a}$ ist $\bar{x} \cap \bar{a} \neq 0$ und aus $\bar{X} \leq \bar{G}$ folgt $\bar{x} \subset \bar{a}$. Also ist $\bar{a}_{\beta+1} \cap \bar{x} \neq 0$, und aus $\bar{G}_0 \leq G$ folgern wir $\bar{a}_{\beta+1} \subset \bar{a}$. Somit haben wir $\bar{a}_\alpha \subset \bar{a}$ und daher auch $\bar{u} \subset \bar{a}$ bewiesen.

d) \bar{U} ist die einzige kleinste Überdeckung von (\bar{X}) . Denn ist auch \bar{G} eine kleinste Überdeckung von (\bar{X}) , so gelten die Beziehungen $\bar{G} \geq \bar{U}$, $\bar{U} \geq \bar{G}$, woraus wegen der Antisymmetrie der Beziehung \geq folgt $\bar{G} = \bar{U}$.

3. Es gibt genau eine größte Verfeinerung von (\bar{X}) . Dieselbe ist durch die im Teile a) des folgenden Beweises gegebene Konstruktion bestimmt.

Beweis. a) Es sei $a, b \in G$. Wir sagen, daß sich a mit b verbinden läßt, wenn a in demselben Elemente von jeder Zerlegung $\bar{X} \in (\bar{X})$ von G wie b liegt. Die für je zwei Elemente $a, b \in G$ definierte Beziehung, daß sich a mit b verbinden läßt, ist offenbar reflexiv, symmetrisch und transitiv. Es gibt also eine Zerlegung \bar{V} von G , so daß sich zwei Elemente in G dann und nur dann verbinden lassen, wenn sie in demselben Elemente von \bar{V} liegen. \bar{V} ist die größte Verfeinerung von (\bar{X}) .

b) \bar{V} ist eine Verfeinerung von (\bar{X}) . Zum Beweise betrachten wir eine Zerlegung $\bar{X} \in (\bar{X})$ von G und zwei inzidente Elemente $\bar{v} \in \bar{V}$, $\bar{x} \in \bar{X}$. Nach 1 genügt es zu zeigen, daß $\bar{x} \supset \bar{v}$ ist. Wegen $\bar{v} \cap \bar{x} \neq 0$ gibt es ein $a \in \bar{v} \cap \bar{x}$ und nach der Definition von \bar{V} läßt sich a mit jedem $b \in \bar{v}$ verbinden. Es ist also $b \in \bar{x}$, und daraus folgt $\bar{x} \supset \bar{v}$.

c) \bar{V} ist eine größte Verfeinerung von (\bar{X}) . Zum Beweise betrachten wir eine Verfeinerung \bar{G} von (\bar{X}) und zwei inzidente Elemente $\bar{v} \in \bar{V}$, $\bar{a} \in \bar{G}$. Wieder genügt es, die Beziehung $\bar{v} \supset \bar{a}$ zu beweisen. Wegen $\bar{v} \cap \bar{a} \neq 0$ gibt es ein $a \in \bar{v} \cap \bar{a}$. Es sei $\bar{X} \in (\bar{X})$, $\bar{x} \in \bar{X}$, $a \in \bar{x}$. Da \bar{G} eine Verfeinerung von (\bar{X}) ist, so haben wir $\bar{a} \subset \bar{x}$. Also läßt sich a mit jedem $b \in \bar{a}$ verbinden, und daraus folgt $\bar{v} \supset \bar{a}$.

d) \bar{V} ist die einzige größte Verfeinerung von (\bar{X}) . Der Beweis ist ähnlich wie in 2d).

Nach den Sätzen 2, 3 bildet das durch die Beziehung \geq teilweise geordnete System aller Zerlegungen von G einen vollständigen Verband (complete lattice). Wie wir schon erwähnt haben, hat dieser vollständige Verband das größte Element \bar{G}_{\max} und das kleinste \bar{G}_{\min} .

Wir beweisen noch den folgenden Satz betreffend die kleinste gemeinsame Überdeckung \bar{U} von zwei Zerlegungen \bar{G}_1, \bar{G}_2 von G .

4. Aus $\bar{a}_1 \in \bar{G}_1$, $\bar{a}_2 \in \bar{G}_2$, $s(\bar{a}_1 \sqsubset \bar{G}_2) = s(\bar{a}_2 \sqsubset \bar{G}_1) (\equiv \bar{u})$ folgt $\bar{u} \in \bar{U}$.

Beweis. Nach der Definition von \bar{u} ist $\bar{u} = \Sigma_1 \bar{b}_1 = \Sigma_2 \bar{b}_2$, wobei sich $\Sigma_1 (\Sigma_2)$ auf die mit $\bar{a}_2 (\bar{a}_1)$ inzidenten Elemente $\bar{b}_1 \in \bar{G}_1 (\bar{b}_2 \in \bar{G}_2)$ bezieht. Da jedes Element von $\bar{a}_1 (\bar{a}_2)$ in einem Elemente von $\bar{G}_2 (\bar{G}_1)$ liegt, so ist $\bar{a}_1, \bar{a}_2 \subset \bar{u}$. Also ist $\bar{a}_1 (\bar{a}_2)$ eines der Elemente $\bar{b}_1 (\bar{b}_2)$ und ist daher mit $\bar{a}_2 (\bar{a}_1)$ inzident. Für $\bar{b}_1 \in \bar{G}_1$, $\bar{b}_1 \subset \bar{u}$ stellt also $\{\bar{a}_1, \bar{b}_1\}$ eine Kette in $\{\bar{G}_1, \bar{G}_2\}$ von \bar{a}_1 nach \bar{b}_1 dar. Wir haben nur noch zu zeigen, daß es für $\bar{b}_1 \in \bar{G}_1$, $\bar{b}_1 \not\subset \bar{u}$ keine Kette in $\{\bar{G}_1, \bar{G}_2\}$ von \bar{a}_1 nach \bar{b}_1 gibt. Gibt es eine

solche Kette $\{\bar{a}_1 = \bar{a}_{10}, \bar{a}_{11}, \dots, \bar{a}_{1\alpha} = \bar{b}_1\}$, so gilt für ein geeignetes $(0 \leq \beta \leq \alpha - 1)$: $\bar{a}_{1\beta} \subset \bar{u}$, $\bar{a}_{1,\beta+1} \not\subset \bar{u}$. Nach der Definition einer Kette gibt es ein mit $\bar{a}_{1\beta}$ und $\bar{a}_{1,\beta+1}$ inzidentes Element $\bar{b}_2 \in \bar{G}_2$. Aus $\bar{a}_{1\beta} \subset \bar{u}$ und $\bar{b}_2 \cap \bar{a}_{1\beta} \neq 0$ folgt $\bar{b}_2 \subset \bar{u}$, und wegen $\bar{b}_2 \cap \bar{a}_{1,\beta+1} \neq 0$ ist $\bar{a}_{1,\beta+1} \subset \bar{u}$.

4. Verknüpfte Zerlegungen. Es seien \bar{A}, \bar{C} Zerlegungen in G . Wir nennen die Zerlegungen \bar{A}, \bar{C} *verknüpft* und sagen, daß $\bar{A}(\bar{C})$ mit $\bar{C}(\bar{A})$ verknüpft ist, wenn es zu jedem Elemente $\bar{a} \in \bar{A}$ genau ein mit ihm inzidentes Element $\bar{c} \in \bar{C}$ gibt und umgekehrt zu jedem Elemente $\bar{c} \in \bar{C}$ genau ein mit ihm inzidentes Element $\bar{a} \in \bar{A}$.

Wir werden nun voraussetzen, daß $\bar{A} = \bar{C} \cap \bar{A}$, $\bar{C} = \bar{A} \cap \bar{C}$. Daraus folgt $s\bar{A} \cap s\bar{C} \neq 0$.

1. Die \bar{A}, \bar{C} sind verknüpft, wenn und nur wenn $\bar{A} \cap s\bar{C} = \bar{C} \cap s\bar{A}$ ist.

Beweis. a) Die \bar{A}, \bar{C} seien verknüpft. Man betrachte z. B. ein Element $\bar{a}' \in \bar{A} \cap s\bar{C}$, so daß $\bar{a}' = \bar{a} \cap s\bar{C}$ gilt, wobei \bar{a} ein geeignetes Element in \bar{A} bedeutet. Nach der Voraussetzung gibt es genau ein mit \bar{a} inzidentes Element $\bar{c} \in \bar{C}$ und \bar{a} ist das einzige mit \bar{c} inzidente Element in \bar{A} . Es ist also $\bar{a}' = \bar{a} \cap \bar{c} = \bar{c} \cap s\bar{A} \in \bar{C} \cap s\bar{A}$.

b) Es sei $\bar{A} \cap s\bar{C} = \bar{C} \cap s\bar{A}$. Wir haben vorausgesetzt $\bar{A} = \bar{C} \cap \bar{A}$, d. h. daß jedes Element in \bar{A} mit wenigstens einem Element in \bar{C} inzident ist. Ist ein Element $\bar{a} \in \bar{A}$ mit $\bar{c}_1, \bar{c}_2 \in \bar{C}$ inzident, so ist $\bar{a} \cap (\bar{c}_1 \vee \bar{c}_2) \subset \bar{a} \cap s\bar{C} \subset \bar{C} \cap s\bar{A}$ und es gibt ein Element $\bar{c} \in \bar{C}$, für das die Beziehung $\bar{a} \cap (\bar{c}_1 \vee \bar{c}_2) \subset s\bar{A} \cap \bar{c}$ besteht. Daraus folgt, da die Elemente in \bar{C} paarweise fremd sind, $\bar{c}_1 = \bar{c}_2 = \bar{c}$.

Wir betrachten nun eine gemeinsame Überdeckung \bar{B} der beiden Zerlegungen $\bar{A} \cap s\bar{C}$, $\bar{C} \cap s\bar{A}$ der Menge $s\bar{A} \cap s\bar{C}$ und definieren die Zerlegung $\bar{A}(\bar{C})$ von $\bar{A}(\bar{C})$ in folgender Weise: Jedes Element in $\bar{A}(\bar{C})$ ist die Menge aller Elemente in $\bar{A}(\bar{C})$, die mit demselben Elemente in \bar{B} inzident sind. Es sei $\hat{A}(\hat{C})$ die durch $\bar{A}(\bar{C})$ erzwungene Überdeckung von $\bar{A}(\bar{C})$; es ist also $\Sigma \hat{a} \in \hat{A}(\Sigma \hat{c} \in \hat{C})$, wenn und nur wenn $\Sigma(\bar{a} \cap s\bar{C}) \in \bar{B}$ ($\Sigma(\bar{c} \cap s\bar{A}) \in \bar{B}$) gilt.

2. Die \hat{A}, \hat{C} sind verknüpft und es gilt $\hat{A} \cap \hat{C} = \bar{B}$.

Beweis. Aus $\bar{A} = \bar{C} \cap \bar{A}$ folgt $\hat{A} = \hat{C} \cap \hat{A}$ und ähnlich $\hat{C} = \hat{A} \cap \hat{C}$. Es genügt zu zeigen, daß

$$(1) \quad \hat{A} \cap s\hat{C} = \hat{C} \cap s\hat{A} = \bar{B}$$

gilt. Denn zu jedem $\hat{a}' \in \hat{A} \cap s\hat{C}$ gibt es ein $\hat{a} \in \hat{A}$, so daß $\hat{a}' = \hat{a} \cap s\hat{C} = \hat{a} \cap \Sigma \hat{c}$, wobei sich Σ auf alle in \hat{C} vorkommenden Elemente \hat{c} bezieht. Ist (1) erfüllt, so sind nach 1 die \hat{A}, \hat{C} verknüpft, und daher gibt es ein

einziges mit \bar{a} inzidentes Element $\bar{c} \in \bar{C}$. Also ist $a' = \bar{a} \cap \bar{c} \in \bar{A} \cap \bar{C}$. Daraus folgt $\bar{A} \cap s\bar{C} \subset \bar{A} \cap \bar{C}$ und ähnlich \supset und daher $\bar{A} \cap s\bar{C} = \bar{A} \cap \bar{C}$. Es sei also $(\Sigma_1 \bar{a}) \cap s\bar{C} \in \bar{A} \cap s\bar{C}$, wobei $\Sigma_1 \bar{a} \in \bar{A}$. Wir haben $(\Sigma_1 \bar{a}) \cap s\bar{C} = \Sigma_1 (\bar{a} \cap s\bar{C}) \in \bar{B}$. Da \bar{B} zugleich eine Überdeckung der Zerlegung $\bar{C} \cap s\bar{A}$ von $s\bar{A} \cap s\bar{C}$ ist, so ist $\Sigma_1 (\bar{a} \cap s\bar{C}) = \Sigma_2 (\bar{c} \cap s\bar{A})$, wobei die \bar{c} in \bar{C} sind. Daraus folgt $\Sigma_2 \bar{c} \in \bar{C}$, also $\Sigma_2 (\bar{c} \cap s\bar{A}) = (\Sigma_2 \bar{c}) \cap s\bar{A} \in \bar{C} \cap s\bar{A}$. Also ist $(\Sigma_1 \bar{a}) \cap s\bar{C} \in \bar{C} \cap s\bar{A}$.

5. Adjungierte Zerlegungen. Wir betrachten Zerlegungen \bar{A}, \bar{C} in G und setzen voraus, daß $B \in \bar{A}, D \in \bar{C}$, wobei $B \cap D \neq 0$. Wir setzen $A = s\bar{A}, C = s\bar{C}$. Aus der Voraussetzung $B \cap D \neq 0$ folgt $B \in D \subset \bar{A}, D \in B \subset \bar{C}$ und wegen $B \subset A, D \subset C$ ist $0 \neq B \cap D \subset B \cap C, D \cap A$. Also sind

$$(1) \quad D \subset \bar{A} \cap C, \quad B \subset \bar{C} \cap A$$

Zerlegungen in G . Wenn für diese Zerlegungen die folgende Gleichheit besteht

$$(2) \quad s(D \subset \bar{A} \cap C) = s(B \subset \bar{C} \cap A),$$

so sagen wir, daß die Zerlegungen \bar{A}, \bar{C} in bezug auf B, D adjungiert sind, oder, daß die Zerlegung $\bar{A} (\bar{C})$ zur $\bar{C} (\bar{A})$ in bezug auf B, D adjungiert ist. Aus § 1 2° folgt, daß man eine äquivalente Definition erhält, wenn man die Symbole (1) durch

$$(1') \quad (D \cap A) \subset (\bar{A} \cap C), \quad (B \cap C) \subset (\bar{C} \cap A)$$

ersetzt, wenn also die folgende Gleichheit gilt:

$$(2') \quad s((D \cap A) \subset (\bar{A} \cap C)) = s((B \cap C) \subset (\bar{C} \cap A)).$$

Die Gleichheit (2) besteht z. B. dann, wenn \bar{A} die größte oder die kleinste Zerlegung von A ist.

Wir setzen nun voraus, daß \bar{A}, \bar{C} in bezug auf B, D adjungiert sind. Dann sind

$$\begin{aligned} \bar{A}_1 &= C \subset \bar{A}, & \bar{C}_1 &= A \subset \bar{C}, \\ \bar{A}_2 &= D \subset \bar{A}, & \bar{C}_2 &= B \subset \bar{C} \end{aligned}$$

Zerlegungen in G . Wir setzen $A_1 = s\bar{A}_1$, usw., und erhalten die folgenden Beziehungen:

$$\begin{aligned} \bar{A} \supset \bar{A}_1 \supset \bar{A}_2 \supset \{B\}, & \quad \bar{C} \supset \bar{C}_1 \supset \bar{C}_2 \supset \{D\}, \\ A \supset A_1 \supset A_2 \supset B, & \quad C \supset C_1 \supset C_2 \supset D. \end{aligned}$$

1. Es gibt verknüpfte Überdeckungen \bar{A}, \bar{C} von \bar{A}_1, \bar{C}_1 , wobei $A_2 \in \bar{A}, C_2 \in \bar{C}$ gilt. Dieselben sind durch die im Teile a) des folgenden Beweises gegebene Konstruktion bestimmt. Die Mengen A_2, C_2 sind inzident.

Beweis. a) Jedes Element in $\bar{A}_1 (\bar{C}_1)$ ist in $\bar{A} (\bar{C})$ enthalten und ist mit $C (A)$, also auch mit einem mit $A (C)$ inzidenten Elemente in $\bar{C} (\bar{A})$ inzident; dieses Element in $\bar{C} (\bar{A})$ ist also in $\bar{C}_1 (\bar{A}_1)$. Es ist also

$$\bar{A}_1 = C_1 \sqsubset \bar{A}_1, \quad \bar{C}_1 = \bar{A}_1 \sqsubset \bar{C}_1.$$

Weiter ist nach 2 5 2

$$A_1 \cap \bar{C}_1 = A_1 \cap \bar{C} = A \cap \bar{C},$$

$$C_1 \cap \bar{A}_1 = C_1 \cap \bar{A} = C \cap \bar{A},$$

so daß $A_1 \cap \bar{C}_1, C_1 \cap \bar{A}_1$ Zerlegungen auf $A \cap C$ sind. Wir bezeichnen mit \bar{U} ihre kleinste gemeinsame Überdeckung. Nach 4 2 gibt es verknüpfte Überdeckungen \bar{A}, \bar{C} von \bar{A}_1, \bar{C}_1 , für die $\bar{A} \cap \bar{C} = \bar{U}$ gilt. Jedes Element in $\bar{A} (\bar{C})$ ist die Summe aller mit demselben Elemente in \bar{U} inzidenten Elemente in $\bar{A}_1 (\bar{C}_1)$. \bar{A}, \bar{C} sind die erwähnten Überdeckungen.

b) Es ist $A_2 \in \bar{A}, C_2 \in C$. In der Tat, aus den Beziehungen $B \in \bar{A}, D \in \bar{C}$ folgt

$$C \cap B \in C \cap \bar{A}, \quad A \cap D \in A \cap \bar{C}$$

und da die \bar{A}, \bar{C} in bezug auf B, D adjungiert sind, so gilt (2'). Es gilt also nach 3 4

$$\bar{u} \in \bar{U},$$

wobei \bar{u} die beiderseits in (2') stehende Menge bedeutet. Die einzelnen Elemente in $\bar{A} (\bar{C})$ sind die Summen aller je mit einem Elemente in \bar{U} inzidenten Elemente in $\bar{A}_1 (\bar{C}_1)$. Zur Feststellung der Beziehung $A_2 \in \bar{A} (C_2 \in \bar{C})$ genügt es also zu beweisen, daß $A_2 (C_2)$ die Summe aller mit \bar{u} inzidenten Elemente in $\bar{A}_1 (\bar{C}_1)$ ist. Nun ist aber

$$\bar{u} = \kappa(D \sqsubset \bar{A} \cap C) = \kappa(\bar{A}_2 \cap C) = A_2 \cap C.$$

Ein Element in \bar{A}_1 ist zugleich in \bar{A} enthalten und ist inzident mit C ; es ist in gleicher Zeit in \bar{A}_2 , wenn und nur wenn es sogar mit D , also auch mit $A_2 \cap C = \bar{u}$ inzident ist. Es sind also mit \bar{u} genau diejenigen Elemente in \bar{A}_1 inzident, die in \bar{A}_2 enthalten sind; ihre Summe ist also A_2 . Ähnlich folgt aus den Gleichungen

$$\bar{u} = \kappa(B \sqsubset \bar{C} \cap A) = \kappa(\bar{C}_2 \cap A) = C_2 \cap A,$$

daß die Summe der mit \bar{u} inzidenten Elemente in \bar{C}_1 die Menge C_2 ist.

c) Aus $0 \neq B \cap D \subset A_2 \cap C_2$ folgt $A_2 \cap C_2 \neq 0$.

II. Grundbegriffe der Gruppoidentheorie.

6. Gruppoid. Wir bezeichnen nach wie vor mit G eine nicht leere Menge. Den Inbegriff von G und einer Multiplikation \mathcal{M} in G nennen wir *Gruppoid*¹⁾. $G(\mathcal{M})$ ist das *Feld* (die *Multiplikation*) des Gruppoides. Gruppoiden bezeichnen wir in der Regel mit denselben jedoch deutschen Buchstaben wie ihre Felder, z. B. \mathfrak{G} , und übertragen auf dieselben die für ihre Felder definierten mengentheoretischen Begriffe und Symbole. Wir sprechen also z. B. von Elementen, Untermengen, Zerlegungen in bzw. auf Gruppoiden und schreiben $a \in \mathfrak{G}$, $A \subset \mathfrak{G}$, usw. Das Produkt von a und b , wobei $a, b \in \mathfrak{G}$, bezeichnen wir mit $a \cdot b$ oder kürzer mit ab . Es ist zu beachten, daß im allgemeinen über die Multiplikation *keine* zusätzlichen Voraussetzungen (wie etwa die Gültigkeit des Assoziativgesetzes usw.) gemacht werden.

Die folgenden Grundbegriffe der Gruppoidentheorie dürfen wir wohl als bekannt voraussetzen:

1) Den Begriff des Produktes AB einer Untermenge $A \subset \mathfrak{G}$ und einer weiteren Untermenge $B \subset \mathfrak{G}$. Ist einer der beiden Faktoren A, B leer, so wollen wir unter dem Symbole AB die leere Menge verstehen.

2) Den Begriff eines Untergruppoides \mathfrak{A} in \mathfrak{G} und eines Obergruppoides \mathfrak{G} über \mathfrak{A} . Bezeichnung: $\mathfrak{A} \subset \mathfrak{G}$, $\mathfrak{G} \supset \mathfrak{A}$. Wenn $\mathfrak{A} \subset \mathfrak{G}$, $\mathfrak{A} \neq \mathfrak{G}$, so nennen wir \mathfrak{A} *echtes Untergruppoid* in \mathfrak{G} , und ähnlich definieren wir ein *echtes Obergruppoid*. Für jedes $\mathfrak{A} \subset \mathfrak{G}$ besteht die Beziehung $\mathfrak{A}A \subset \mathfrak{A}$; gilt umgekehrt diese Beziehung für eine nicht leere Untermenge $A \subset \mathfrak{G}$, so bildet dieselbe samt der durch \mathcal{M} bestimmten Multiplikation in A ein Untergruppoid in \mathfrak{G} .

3) Der Begriff des Durchschnittes und der Vereinigung von zwei Untergruppoiden $\mathfrak{A}, \mathfrak{B}$. Bezeichnungen: $\mathfrak{A} \cap \mathfrak{B}$, $\mathfrak{A} \cup \mathfrak{B}$.

Im folgenden bedeutet $\mathfrak{G} = (G, \mathcal{M})$ ein Gruppoid.

7. Erzeugende Zerlegungen in Gruppoiden. Eine Zerlegung \bar{A} in \mathfrak{G} nennen wir *erzeugende Zerlegung* in \mathfrak{G} , wenn es zu jedem geordneten Paare \bar{a}, \bar{b} von Elementen in \bar{A} ein Element $\bar{c} \in \bar{A}$ gibt, so daß die folgende Beziehung besteht: $\bar{a}\bar{b} \subset \bar{c}$. Als Beispiel von erzeugenden Zerlegungen in \mathfrak{G} führen wir die beiden extremen Zerlegungen \bar{G}_{\max} , \bar{G}_{\min} von \mathfrak{G} an.

¹⁾ Der Name „Gruppoid“ für diesen Begriff findet sich in der Arbeit *B. A. Hausmann and Oystein Ore: Theory of quasigroups* (Amer. J. Math., Vol. LIX, 1937). Im engeren Sinne kommt er vor bei *Garrett Birkhoff: Rings of sets* (Duke Math. J., Vol. 3, 1937, p. 444). Siehe auch *H. Brandt: Über eine Verallgemeinerung des Gruppenbegriffes* (Math. Annalen, Bd. 96, 1927).

Es seien \bar{A}, \bar{C} erzeugende Zerlegungen in \mathfrak{G} .

1. Es ist $s\bar{A} \cdot s\bar{A} \subset s\bar{A}$.

Denn für $a; b \in s\bar{A}$ gibt es Elemente $\bar{a}, \bar{b}, \bar{c} \in \bar{A}$, so daß die Beziehungen $a \in \bar{a}, b \in \bar{b}, \bar{a}\bar{b} \subset \bar{c}$ bestehen. Also ist $ab \in \bar{a}\bar{b} \subset \bar{c} \subset s\bar{A}$.

Nach diesem Satze ist der Inbegriff der Untermenge $s\bar{A}$ und der durch \mathcal{M} bestimmten Multiplikation in $s\bar{A}$ ein Untergruppoid in \mathfrak{G} .

2. Wenn $s\bar{A} \cap s\bar{C} \neq 0$ ist, so sind $\bar{A} \cap \bar{C}$ und $\bar{A} \subset \bar{C}$ erzeugende Zerlegungen in \mathfrak{G} .

Beweis. Unter der obigen Voraussetzung sind $\bar{A} \cap \bar{C}$ und $\bar{A} \subset \bar{C}$ Zerlegungen in \mathfrak{G} . a) Wir betrachten zwei Elemente $\bar{x}, \bar{y} \in \bar{A} \cap \bar{C}$. Nach der Definition von $\bar{A} \cap \bar{C}$ gibt es Elemente $\bar{a}_1, \bar{a}_2 \in \bar{A}; \bar{c}_1, \bar{c}_2 \in \bar{C}$, so daß $\bar{x} = \bar{a}_1 \cap \bar{c}_1, \bar{y} = \bar{a}_2 \cap \bar{c}_2$. Da $\bar{A} (\bar{C})$ eine erzeugende Zerlegung ist, so gibt es ein $\bar{a} \in \bar{A} (\bar{c} \in \bar{C})$, so daß $\bar{a}_1 \bar{a}_2 \subset \bar{a} (\bar{c}_1 \bar{c}_2 \subset \bar{c})$. Nun ist aber $\bar{x} \bar{y} \subset \bar{a}_1 \bar{a}_2 \cap \bar{c}_1 \bar{c}_2 \subset \bar{a} \cap \bar{c} \in \bar{A} \cap \bar{C}$. b) Wegen $\bar{A} \subset \bar{C} = s\bar{A} \subset \bar{C}$ genügt es zu beweisen, daß $s\bar{A} \subset \bar{C}$ erzeugend ist. Es seien $\bar{c}_1, \bar{c}_2 \in s\bar{A} \subset \bar{C}$. Da \bar{C} erzeugend ist, so gibt es ein $\bar{c} \in \bar{C}$, so daß $\bar{c}_1 \bar{c}_2 \subset \bar{c}$. Wir wählen ein $x \in s\bar{A} \cap \bar{c}_1$ und ein $y \in s\bar{A} \cap \bar{c}_2$. Es ist $xy \in s\bar{A} \cdot s\bar{A} \cap \bar{c}_1 \bar{c}_2 \subset s\bar{A} \cap \bar{c}$, woraus $s\bar{A} \cap \bar{c} \neq 0$ folgt. Also ist $\bar{c} \in s\bar{A} \subset \bar{C}$.

Wir setzen nun $\bar{A} = \bar{C} \cap \bar{A}, \bar{C} = \bar{A} \cap \bar{C}$ voraus und betrachten eine gemeinsame Überdeckung \bar{B} der beiden auf der Menge $s\bar{A} \cap s\bar{C}$ liegenden Zerlegungen $\bar{A} \cap s\bar{C}, \bar{C} \cap s\bar{A}$ und definieren die Überdeckungen \bar{A}, \bar{C} von \bar{A}, \bar{C} in derselben Weise wie in 4.2.

3. Wenn \bar{B} erzeugend ist, so sind es auch \bar{A}, \bar{C} .

Beweis. Es sei $\Sigma_1 \bar{a}_1, \Sigma_2 \bar{a}_2 \in \bar{A}$; also sind die \bar{a}_1, \bar{a}_2 Elemente in \bar{A} und $\Sigma_1 (\bar{a}_1 \cap s\bar{C}), \Sigma_2 (\bar{a}_2 \cap s\bar{C})$ sind Elemente in \bar{B} . Da \bar{B} erzeugend ist, so gibt es zu jedem Produkte $\bar{a}_1 \bar{a}_2$ ein $\bar{a}_{12} \in \bar{A}$, so daß $\bar{a}_1 \bar{a}_2 \subset \bar{a}_{12}$, woraus $(\bar{a}_1 \cap s\bar{C}) (\bar{a}_2 \cap s\bar{C}) \subset \bar{a}_{12} \cap s\bar{C}$ folgt. Da \bar{B} erzeugend ist, so gibt es ein $\Sigma_3 (\bar{a}_3 \cap s\bar{C}) \in \bar{B}$, so daß $\Sigma_1 (\bar{a}_1 \cap s\bar{C}) \Sigma_2 (\bar{a}_2 \cap s\bar{C}) = \Sigma_1 \Sigma_2 (\bar{a}_1 \cap s\bar{C}) (\bar{a}_2 \cap s\bar{C}) \subset \Sigma_3 (\bar{a}_3 \cap s\bar{C})$ ist. Dabei sind die \bar{a}_3 Elemente in \bar{A} und $\Sigma_3 \bar{a}_3 \in \bar{A}$. Für jedes $\bar{a}_1 (\bar{a}_2)$, auf das sich $\Sigma_1 (\Sigma_2)$ bezieht, haben wir also die Beziehung $(\bar{a}_1 \cap s\bar{C}) (\bar{a}_2 \cap s\bar{C}) \subset (\bar{a}_{12} \cap s\bar{C}) \cap \Sigma_3 (\bar{a}_3 \cap s\bar{C})$. Die $\bar{a}_{12} \cap s\bar{C}, \bar{a}_3 \cap s\bar{C}$ sind aber Elemente in der Zerlegung $\bar{A} \cap s\bar{C}$ von $s\bar{A} \cap s\bar{C}$. Also gilt für ein geeignetes hinter dem Zeichen Σ_3 stehendes \bar{a}_3 die Beziehung $\bar{a}_{12} \cap s\bar{C} = \bar{a}_3 \cap s\bar{C}$, woraus $\bar{a}_{12} = \bar{a}_3$ folgt. Wir haben also schließlich $\Sigma_1 \bar{a}_1 \Sigma_2 \bar{a}_2 \subset \Sigma_1 \Sigma_2 \bar{a}_{12} \subset \Sigma_3 \bar{a}_3 \in \bar{A}$, womit der Satz bewiesen ist.

Es sei nun (\bar{X}) ein nicht leeres System von erzeugenden Zerlegungen von \mathfrak{G} .

4. Die kleinste Überdeckung \bar{U} von (\bar{X}) ist erzeugend.

Beweis. Es sei $\bar{u}, \bar{v} \in \bar{U}$. Wir haben zu zeigen, daß es ein $\bar{w} \in \bar{U}$ gibt, so daß $\bar{u} \bar{v} \subset \bar{w}$. Wir betrachten eine Zerlegung $\bar{G}_0 \in (\bar{X})$ von G und beliebige Elemente $\bar{a}_0, \bar{b}_0 \in \bar{G}_0$; $\bar{a}_0 \subset \bar{u}$, $\bar{b}_0 \subset \bar{v}$. Da G_0 erzeugend ist, so gibt es ein $\bar{c}_0 \in \bar{G}_0$, so daß $\bar{a}_0 \bar{b}_0 \subset \bar{c}_0$. Wir bezeichnen mit \bar{w} dasjenige Element in \bar{U} , für welches $\bar{c}_0 \subset \bar{w}$ gilt.

Es sei $\bar{a}_\alpha, \bar{b}_\beta \in \bar{G}_0$, $\bar{a}_\alpha \subset \bar{u}$, $\bar{b}_\beta \subset \bar{v}$. Es gibt ein $\bar{c}_\gamma \in \bar{G}_0$, für welches die Beziehung $\bar{a}_\alpha \bar{b}_\beta \subset \bar{c}_\gamma$ besteht, und wir haben zu zeigen, daß $\bar{c}_\gamma \subset \bar{w}$ gilt. Nach der Konstruktion von \bar{U} gibt es eine Kette in (\bar{X}) von \bar{a}_0 nach \bar{a}_α : $\{\bar{a}_0, \dots, \bar{a}_\alpha\}$ und eine Kette in (\bar{X}) von \bar{b}_0 nach \bar{b}_β : $\{\bar{b}_0, \dots, \bar{b}_\beta\}$. Wir können $\beta = \alpha$ voraussetzen, denn z. B. im Falle $\beta < \alpha$ genügt es, zu der zweiten Kette $\alpha - \beta$ gleiche Elemente \bar{b}_β zuzufügen, um diese Voraussetzung zu erfüllen. Nun gibt es Elemente $\bar{c}_0, \dots, \bar{c}_{2\alpha} \in \bar{G}_0$, so daß für $\mu = 0, \dots, \alpha$ die Beziehungen $\bar{c}_{2\mu} \supset \bar{a}_\mu \bar{b}_\mu$, $\bar{c}_{2\mu+1} \supset \bar{a}_{\mu+1} \bar{b}_\mu$ ($\bar{c}_{2\alpha+1} = \bar{a}_{\alpha+1} = 0$, $\bar{c}_{2\alpha} = \bar{c}_\gamma$) bestehen, und es genügt offenbar zu beweisen, daß $\{\bar{c}_0, \dots, \bar{c}_{2\alpha}\}$ eine Kette in (\bar{X}) von \bar{c}_0 nach $\bar{c}_{2\alpha}$ ist. Zu diesem Zwecke werden wir zeigen, daß zu je zwei einander folgenden Elementen \bar{c}_r, \bar{c}_{r+1} ($r = 0, \dots, 2\alpha - 1$) ein mit beiden inzidentes Element in einer geeigneten Zerlegung $\bar{G} \in (\bar{X})$ von G existiert. Es sei z. B. $r = 2\mu$ eine gerade Zahl. Nach der Voraussetzung gibt es in einer geeigneten Zerlegung $\bar{G} \in (\bar{X})$ ein mit $\bar{a}_\mu, \bar{a}_{\mu+1}$ inzidentes Element \bar{x} ; es existieren also $a', b' \in G$, so daß $a' \in \bar{a}_\mu \cap \bar{x}$, $b' \in \bar{a}_{\mu+1} \cap \bar{x}$. Wir wählen ein $a'' \in \bar{b}_\mu$. Dann gibt es ein $\bar{y} \in \bar{G}$, so daß $a'' \in \bar{b}_\mu \cap \bar{y}$. Da \bar{G} erzeugend ist, so gibt es ein $\bar{z} \in \bar{G}$, für welches $\bar{x} \bar{y} \subset \bar{z}$ gilt. Nun haben wir $a' a'' \in \bar{x} \bar{y} \cap \bar{a}_\mu \bar{b}_\mu \subset \bar{z} \cap \bar{c}_r$, und desgleichen $b' a'' \in \bar{x} \bar{y} \cap \bar{a}_{\mu+1} \bar{b}_\mu \subset \bar{z} \cap \bar{c}_{r+1}$, woraus folgt, daß $a' a''$ ($b' a''$) sowohl in \bar{z} als auch in \bar{c}_r (\bar{c}_{r+1}) enthalten ist.

5. Die größte Verfeinerung \bar{V} von (\bar{X}) ist ebenfalls eine erzeugende.

Beweis. Es sei $\bar{u}, \bar{v} \in \bar{V}$. Wir haben zu zeigen, daß es ein $\bar{w} \in \bar{V}$ gibt, so daß $\bar{u} \bar{v} \subset \bar{w}$. Wir betrachten beliebige Elemente $a \in \bar{u}$, $b \in \bar{v}$ von G und bezeichnen mit \bar{w} dasjenige Element in \bar{V} , welches ab enthält. Weiter betrachten wir beliebige Elemente $x \in \bar{u}$, $y \in \bar{v}$ in G ; a (b) läßt sich also mit x (y) verbinden. Offenbar genügt es zu zeigen, daß ab sich mit xy verbinden läßt. Es sei also $\bar{G} \in (\bar{X})$, $\bar{c} \in \bar{G}$, $ab \in \bar{c}$. Wir haben zu zeigen, daß $xy \in \bar{c}$. Wir bezeichnen mit \bar{a} (\bar{b}) dasjenige Element in \bar{G} , welches a (b) enthält. Nach der Voraussetzung haben wir $x \in \bar{a}$, $y \in \bar{b}$, also $xy \in \bar{a} \bar{b}$. Da \bar{G} erzeugend ist, so gibt es ein $\bar{c}' \in \bar{G}$, für welches $\bar{a} \bar{b} \subset \bar{c}'$ gilt. Nun ist aber $ab \in \bar{c} \cap \bar{a} \bar{b} \subset \bar{c} \cap \bar{c}'$, woraus $\bar{c}' = \bar{c}$ folgt. Also ist $xy \in \bar{c}$.

Nach den beiden letzten Sätzen ist die durch die Beziehung \geq teilweise geordnete Menge aller erzeugenden Zerlegungen von G eine voll-

ständiger Verband. Da die \bar{G}_{\max} und \bar{G}_{\min} erzeugende Zerlegungen von \mathfrak{G} sind, so hat dieser vollständige Verband das größte Element \bar{G}_{\max} und das kleinste \bar{G}_{\min} .

8. Faktoroid. Es sei \bar{A} eine erzeugende Zerlegung in \mathfrak{G} . Wir definieren eine Multiplikation \mathcal{N} in \bar{A} in folgender Weise: Für $\bar{a}, \bar{b} \in \bar{A}$ ist das Produkt von \bar{a} und \bar{b} dasjenige Element $\bar{c} \in \bar{A}$, für welches $\bar{a} \bar{b} \subset \bar{c}$ gilt. Das Gruppoid $\bar{\mathfrak{A}} = (\bar{A}, \mathcal{N})$ nennen wir *Faktoroid* in \mathfrak{G} und im Falle, daß \bar{A} auf \mathfrak{G} liegt, *Faktoroid* von \mathfrak{G} oder *auf* \mathfrak{G} . Wir schreiben $\bar{a} \cdot \bar{b} = \bar{c}$ und haben daher $\bar{a} \bar{b} \subset \bar{a} \cdot \bar{b} \in \bar{\mathfrak{A}}$. Durch \bar{A} ist also das entsprechende Faktoroid $\bar{\mathfrak{A}}$ eindeutig bestimmt. Das Untergruppoid in \mathfrak{G} , dessen Feld $s\bar{A}$ ist, bezeichnen wir mit $s\bar{\mathfrak{A}}$. Enthält $\bar{\mathfrak{A}}$ das Feld B eines Untergruppoides $\mathfrak{B} \subset \mathfrak{G}$ als Element, so sagen wir, $\bar{\mathfrak{A}}$ enthalte das Untergruppoid \mathfrak{B} als Element und schreiben $\mathfrak{B} \in \bar{\mathfrak{A}}$ oder $\bar{\mathfrak{A}} \ni \mathfrak{B}$. Als Beispiele von Faktoroiden in \mathfrak{G} führen wir an das größte Faktoroid $\bar{\mathfrak{G}}_{\max}$ und das kleinste Faktoroid $\bar{\mathfrak{G}}_{\min}$ von \mathfrak{G} .

Es seien $\bar{\mathfrak{A}}, \bar{\mathfrak{C}}$ Faktoroiden in \mathfrak{G} . Entsprechend dem Satze 7'2 können wir im Falle $s\bar{A} \cap s\bar{C} \neq 0$ die Begriffe *der Durchdringung* von $\bar{\mathfrak{A}}$ und $\bar{\mathfrak{C}}$ und *der Hülle* von $\bar{\mathfrak{A}}$ in $\bar{\mathfrak{C}}$ definieren; Bezeichnungen: $\bar{\mathfrak{A}} \cap \bar{\mathfrak{C}}$ oder $\bar{\mathfrak{C}} \cap \bar{\mathfrak{A}}$ und $\bar{\mathfrak{A}} \subset \bar{\mathfrak{C}}$ oder $\bar{\mathfrak{C}} \supset \bar{\mathfrak{A}}$. Ebenso offenbar sind die Begriffe *der Durchdringung* $\mathfrak{B} \cap \bar{\mathfrak{A}}$ oder $\bar{\mathfrak{A}} \cap \mathfrak{B}$ von $\bar{\mathfrak{A}}$ mit einem Untergruppoid $\mathfrak{B} \subset \mathfrak{G}$ und *der Hülle* $\mathfrak{B} \subset \bar{\mathfrak{A}}$ oder $\bar{\mathfrak{A}} \supset \mathfrak{B}$ von \mathfrak{B} in $\bar{\mathfrak{A}}$ und die Begriffe von *verknüpften* und *adjungierten* Faktoroiden.

Es seien $\bar{\mathfrak{A}} \ni \mathfrak{B}$, $\bar{\mathfrak{C}} \ni \mathfrak{D}$ adjungierte Faktoroiden in bezug auf \mathfrak{B} , \mathfrak{D} . Wir setzen $\bar{\mathfrak{A}} = s\bar{\mathfrak{A}}$, usw. Dann sind nach 7'2

$$\begin{aligned} \bar{\mathfrak{A}}_1 &= \mathfrak{C} \subset \bar{\mathfrak{A}}, & \bar{\mathfrak{C}}_1 &= \mathfrak{A} \subset \bar{\mathfrak{C}}, \\ \bar{\mathfrak{A}}_2 &= \mathfrak{D} \subset \bar{\mathfrak{A}}, & \bar{\mathfrak{C}}_2 &= \mathfrak{B} \subset \bar{\mathfrak{C}} \end{aligned}$$

Faktoroiden in \mathfrak{G} .

Aus 5'1 (mit Benützung von 7'2'4'3) folgt der folgende Satz:

1. Es gibt verknüpfte Überdeckungen $\bar{\mathfrak{A}}, \bar{\mathfrak{C}}$ von $\bar{\mathfrak{A}}_1, \bar{\mathfrak{C}}_1$, so daß $\bar{\mathfrak{A}}_2 \in \bar{\mathfrak{A}}$, $\bar{\mathfrak{C}}_2 \in \bar{\mathfrak{C}}$; dieselben sind durch die in 5'1 a) gegebene Konstruktion bestimmt. Die Untergruppoiden $\bar{\mathfrak{A}}_2, \bar{\mathfrak{C}}_2$ sind inzident.

Wir betrachten nun Faktoroiden auf \mathfrak{G} :

Es seien $\bar{\mathfrak{G}}_1, \bar{\mathfrak{G}}_2$ Faktoroiden auf \mathfrak{G} . Wir schreiben $\bar{\mathfrak{G}}_1 \geq \bar{\mathfrak{G}}_2$ oder $\bar{\mathfrak{G}}_2 \leq \bar{\mathfrak{G}}_1$, wenn für die Felder \bar{G}_1, \bar{G}_2 von $\bar{\mathfrak{G}}_1, \bar{\mathfrak{G}}_2$ die Beziehung $\bar{G}_1 \geq \bar{G}_2$ besteht; in diesem Falle nennen wir $\bar{\mathfrak{G}}_1(\bar{\mathfrak{G}}_2)$ *Überdeckung* (*Verfeinerung*) von $\bar{\mathfrak{G}}_2(\bar{\mathfrak{G}}_1)$ und sagen, das Faktoroid $\bar{\mathfrak{G}}_1(\bar{\mathfrak{G}}_2)$ sei oder liege auf $\bar{\mathfrak{G}}_2(\bar{\mathfrak{G}}_1)$. $\bar{\mathfrak{G}}_{\max}(\bar{\mathfrak{G}}_2)$ ist die größte (kleinste) Überdeckung von $\bar{\mathfrak{G}}_2$; $\bar{\mathfrak{G}}_1(\bar{\mathfrak{G}}_{\min})$ ist die

größte (kleinste) Verfeinerung von $\overline{\mathfrak{G}}_1$. Durch jede Überdeckung $\overline{\mathfrak{G}}_1$ von $\overline{\mathfrak{G}}_2$ ist eine (und zwar die Überdeckung $\overline{\mathfrak{G}}_1$ erzwingende) Zerlegung $\overline{\mathfrak{G}}_2$ von $\overline{\mathfrak{G}}_2$ eindeutig bestimmt. Umgekehrt erzwingt jede Zerlegung $\overline{\mathfrak{G}}_2$ von $\overline{\mathfrak{G}}_2$ eine bestimmte Überdeckung $\overline{\mathfrak{G}}_1$ der erzeugenden Zerlegung $\overline{\mathfrak{G}}_2$ von \mathfrak{G} . Wenn $\overline{\mathfrak{G}}_1$ erzeugend ist, so ist das Faktoroid $\overline{\mathfrak{G}}_1$, dessen Feld $\overline{\mathfrak{G}}_1$ ist, die durch $\overline{\mathfrak{G}}_2$ erzwungene Überdeckung von $\overline{\mathfrak{G}}_2$. Wann dieser Fall eintritt, darüber belehrt der folgende Satz.

2. Es sei $\overline{\mathfrak{G}}_2$ ein Faktoroid auf \mathfrak{G} und $\overline{\mathfrak{G}}_1$ die durch eine Zerlegung $\overline{\mathfrak{G}}_2$ von $\overline{\mathfrak{G}}_2$ erzwungene Überdeckung der erzeugenden Zerlegung $\overline{\mathfrak{G}}_2$. Die Zerlegung $\overline{\mathfrak{G}}_1$ ist dann und nur dann erzeugend, wenn dieses für die Zerlegung $\overline{\mathfrak{G}}_2$ der Fall ist.

Beweis. a) Wir setzen voraus, daß $\overline{\mathfrak{G}}_2$ erzeugend ist. Es sei $\bar{a}_1, \bar{b}_1 \in \overline{\mathfrak{G}}_1$. Wir haben zu zeigen, daß es ein $\bar{c}_1 \in \overline{\mathfrak{G}}_1$ gibt, so daß $\bar{a}_1 \bar{b}_1 \subset \bar{c}_1$. Nun ist aber $\bar{a}_1 = \Sigma \bar{a}_2, \bar{b}_1 = \Sigma \bar{b}_2$, wobei sich die erste (zweite) Summe auf alle in demselben Elemente $\bar{a}_2 (\bar{b}_2)$ von $\overline{\mathfrak{G}}_2$ liegenden Elemente von $\overline{\mathfrak{G}}_2$ bezieht. Da $\overline{\mathfrak{G}}_2$ erzeugend ist, so gibt es ein $\bar{c}_2 \in \overline{\mathfrak{G}}_2$, so daß $\bar{a}_2 \cdot \bar{b}_2 \subset \bar{c}_2$ gilt. Wir bezeichnen mit \bar{c}_1 die Summe der in \bar{c}_2 liegenden Elemente von $\overline{\mathfrak{G}}_2$ und haben $\bar{c}_1 \in \overline{\mathfrak{G}}_1$. Offenbar gilt für jedes $\bar{a}_2' (\bar{b}_2')$, auf welches sich die erste (zweite) Summe bezieht: $\bar{a}_2' \cdot \bar{b}_2' \in \bar{a}_2 \cdot \bar{b}_2 \subset \bar{c}_2$. Also ist $\bar{a}_1 \bar{b}_1 = \Sigma \Sigma \bar{a}_2 \bar{b}_2 \subset \Sigma \Sigma \bar{a}_2 \cdot \bar{b}_2 \subset \bar{c}_1$.

b) Wir setzen nun voraus, daß $\overline{\mathfrak{G}}_1$ erzeugend ist, und behalten die Bedeutung für $\bar{a}_1, \bar{b}_1; \bar{a}_2, \bar{b}_2; \bar{a}_2', \bar{b}_2'$. Da $\overline{\mathfrak{G}}_1$ erzeugend ist, so gibt es ein $\bar{c}_1 \in \overline{\mathfrak{G}}_1$, für welches $\bar{a}_1 \bar{b}_1 \subset \bar{c}_1$ gilt. Nach der Definition von $\overline{\mathfrak{G}}_1$ gibt es Elemente $\bar{c}_2 \in \overline{\mathfrak{G}}_2$, so daß $\bar{c}_1 = \Sigma \bar{c}_2$, und die Menge dieser Elemente ist ein Element $\bar{c}_2 \in \overline{\mathfrak{G}}_2$. Für $\bar{a}_2 \in \bar{a}_2', \bar{b}_2 \in \bar{b}_2'$ haben wir $\bar{a}_2 \bar{b}_2 \subset \bar{c}_1$ und daher gibt es ein $\bar{c}_2 \in \bar{c}_2$, so daß $\bar{a}_2 \bar{b}_2 \subset \bar{c}_2$. Daraus folgt $\bar{a}_2 \cdot \bar{b}_2 = \bar{c}_2 \in \bar{c}_2$ und schließlich $\bar{a}_2 \cdot \bar{b}_2 \subset \bar{c}_2$.

Es sei $\overline{\mathfrak{G}}$ ein Faktoroid auf \mathfrak{G} . Nach dem Satze 2 ist jede Überdeckung von $\overline{\mathfrak{G}}$ durch eine erzeugende Zerlegung $\overline{\mathfrak{G}}$ von $\overline{\mathfrak{G}}$ erzwungen. Wir sagen auch, daß die Überdeckung durch das entsprechende Faktoroid $\overline{\mathfrak{G}}$ von $\overline{\mathfrak{G}}$ erzwungen ist. Offenbar ist die größte (kleinste) Überdeckung von $\overline{\mathfrak{G}}$ durch das größte (kleinste) Faktoroid von $\overline{\mathfrak{G}}$ erzwungen.

Wir setzen nun voraus, daß $\overline{\mathfrak{G}}$ ein Untergruppoid $\mathfrak{G}_1 \subset \mathfrak{G}$ als Element enthält und bezeichnen mit $\overline{\mathfrak{G}}$ eine Überdeckung von $\overline{\mathfrak{G}}$.

3. Es gibt genau ein durch die Beziehungen $\overline{\mathfrak{G}} \supset \mathfrak{G}_2 \supset \mathfrak{G}_1$ bestimmtes Untergruppoid \mathfrak{G}_2 in \mathfrak{G} . Das Faktoroid $\mathfrak{G}_2 \cap \overline{\mathfrak{G}}$ ist ein Untergruppoid in $\overline{\mathfrak{G}}$.

Beweis. Offenbar gibt es höchstens ein solches Untergruppoid \mathfrak{G}_2 . Wir betrachten das die Überdeckung $\overline{\mathfrak{G}}$ erzwingende Faktoroid $\overline{\mathfrak{G}}$ von $\overline{\mathfrak{G}}$. Das

Feld G_1 von \mathfrak{G}_1 ist in einem Elemente \bar{a} von $\bar{\mathfrak{G}}$ als Element enthalten und aus $G_1 G_1 \subset G_1$ folgt $\bar{a} \cdot \bar{a} \subset \bar{a}$. Wir bezeichnen mit $\bar{\mathfrak{G}}_2$ das Untergruppoid in $\bar{\mathfrak{G}}$, dessen Feld \bar{a} ist, und setzen $\mathfrak{G}_2 = s \bar{\mathfrak{G}}_2$. Offenbar hat das Untergruppoid \mathfrak{G}_2 die obigen Eigenschaften.

9. Relativ einfache und einfache Gruppoide. Es sei $a \in \mathfrak{G}$. Wenn jedes Faktoroid von \mathfrak{G} entweder $\bar{\mathfrak{G}}_{\max}$ ist oder die aus dem einzigen Elemente a bestehende Menge $\{a\}$ als Element enthält, so nennen wir \mathfrak{G} *einfach in bezug auf a* . Wenn \mathfrak{G} in bezug auf eines seiner Elemente einfach ist, so sagen wir, \mathfrak{G} sei *relativ einfach*. Wenn \mathfrak{G} in bezug auf jedes seiner Elemente einfach ist, so nennen wir es *einfach*.

Wenn \mathfrak{G} einfach ist, so sind $\bar{\mathfrak{G}}_{\max}$ und $\bar{\mathfrak{G}}_{\min}$ die einzigen Faktoroide von \mathfrak{G} und umgekehrt.

Beweis. Aus der Definition folgt, daß \mathfrak{G} einfach ist, wenn $\bar{\mathfrak{G}}_{\max}$ und $\bar{\mathfrak{G}}_{\min}$ die einzigen Faktoroide von \mathfrak{G} sind. Wir setzen also voraus, daß \mathfrak{G} einfach ist, und betrachten ein Faktoroid $\bar{\mathfrak{G}}$ von \mathfrak{G} . Wenn $\bar{\mathfrak{G}} \neq \bar{\mathfrak{G}}_{\max}$ ist, so enthält $\bar{\mathfrak{G}}$ für jedes $a \in \bar{\mathfrak{G}}$ die aus dem einzigen Elemente a bestehende Menge $\{a\}$ als Element, woraus $\bar{\mathfrak{G}} = \bar{\mathfrak{G}}_{\min}$ folgt.

Wir betrachten nun relativ einfache und einfache Faktoroide.

Es sei $\bar{\mathfrak{G}}$ ein Faktoroid von \mathfrak{G} und es sei $\mathfrak{G}_1 \in \bar{\mathfrak{G}}$. Wenn $\bar{\mathfrak{G}}$ in bezug auf das Element G_1 , das Feld von \mathfrak{G}_1 , einfach ist, so sagen wir auch, $\bar{\mathfrak{G}}$ sei *einfach in bezug auf \mathfrak{G}_1* .

2. Dann und nur dann, wenn $\bar{\mathfrak{G}}$ in bezug auf \mathfrak{G}_1 einfach ist, ist jede Überdeckung von $\bar{\mathfrak{G}}$ entweder $\bar{\mathfrak{G}}_{\max}$ oder sie enthält \mathfrak{G}_1 als Element.

Beweis. a) $\bar{\mathfrak{G}}$ sei einfach in bezug auf \mathfrak{G}_1 . Wir betrachten eine Überdeckung $\bar{\mathfrak{G}}$ von $\bar{\mathfrak{G}}$ und das diese Überdeckung erzwingende Faktoroid $\bar{\mathfrak{G}}$ von $\bar{\mathfrak{G}}$. Da $\bar{\mathfrak{G}}$ in bezug auf \mathfrak{G}_1 einfach ist, so ist $\bar{\mathfrak{G}}$ entweder das größte Faktoroid von $\bar{\mathfrak{G}}$ oder es enthält die aus dem Felde G_1 von \mathfrak{G}_1 bestehende Menge als Element. Im ersten Falle ist $\bar{\mathfrak{G}}$ das größte Faktoroid von $\bar{\mathfrak{G}}$, im zweiten ist $\bar{\mathfrak{G}} \ni \mathfrak{G}_1$.

b) $\bar{\mathfrak{G}}$ sei nicht einfach in bezug auf \mathfrak{G}_1 . Dann gibt es ein Faktoroid $\bar{\mathfrak{G}}$ von $\bar{\mathfrak{G}}$, das weder das größte Faktoroid von $\bar{\mathfrak{G}}$ ist, noch die Menge G_1 als Element enthält. Es sei $\bar{a} \in \bar{\mathfrak{G}}$, $G_1 \in \bar{a}$. Dann ist G_1 echt in $s \bar{a}$ und $s \bar{a}$ ist echt in G . Die durch $\bar{\mathfrak{G}}$ erzwungene Überdeckung $\bar{\mathfrak{G}}$ von $\bar{\mathfrak{G}}$ enthält $s \bar{a}$ als Element. Also ist $\bar{\mathfrak{G}}$ weder das größte Faktoroid von $\bar{\mathfrak{G}}$, noch enthält es \mathfrak{G}_1 als Element.

Ähnlich beweist man den folgenden Satz:

3. Dann und nur dann, wenn $\bar{\mathfrak{G}}$ einfach ist, ist jede Überdeckung von $\bar{\mathfrak{G}}$ entweder $\bar{\mathfrak{G}}_{\max}$ oder das Faktoroid $\bar{\mathfrak{G}}$ selbst.

10. Die Isomorphiesätze. Den Begriff der *homomorphen Abbildung* eines Gruppoides \mathfrak{G} in und auf ein Gruppoid \mathfrak{G}^* dürfen wir wohl als be-

kannt voraussetzen. Eine homomorphe Abbildung von \mathfrak{G} auf \mathfrak{G}^* heißt auch *Homomorphismus*. Insbesondere setzen wir auch den Begriff einer *isomorphen Abbildung* (*Isomorphismus*) als bekannt voraus. Ein wichtiges Beispiel eines Homomorphismus ist folgendes: Die Abbildung von \mathfrak{G} auf ein Faktoroid $\overline{\mathfrak{G}}$ von \mathfrak{G} , in der jedes Element $a \in \mathfrak{G}$ auf das dieses Element a enthaltende Element $\bar{a} \in \overline{\mathfrak{G}}$ abgebildet wird. Ein wichtiges Beispiel eines Isomorphismus ist folgendes: Ist $\overline{\mathfrak{G}}$ ein Faktoroid von \mathfrak{G} und \mathfrak{G} die durch ein Faktoroid $\overline{\mathfrak{G}}$ von \mathfrak{G} erzwungene Überdeckung von $\overline{\mathfrak{G}}$, so ist die Abbildung von $\overline{\mathfrak{G}}$ auf \mathfrak{G} , in der jedes Element $\bar{a} \in \overline{\mathfrak{G}}$ auf die Summe $\dot{a} \in \mathfrak{G}$ der in \bar{a} liegenden Elemente $\dot{a} \in \overline{\mathfrak{G}}$ abgebildet wird, ein Isomorphismus.

Es sei d eine homomorphe Abbildung von \mathfrak{G} in \mathfrak{G}^* . Für $a \in \mathfrak{G}$ ($0 \neq A \subset \mathfrak{G}$) bezeichnen wir mit da (dA) das Bild von a (die Menge der Bilder der Elemente in A) in d . Ist $\bar{A} = \{\bar{a}, \bar{b}, \dots\}$ eine Zerlegung in \mathfrak{G} , so bedeutet $d\bar{A}$ die Zerlegung $\{d\bar{a}, d\bar{b}, \dots\}$ in \mathfrak{G}^* . Ist also d ein Homomorphismus und \bar{A} eine Zerlegung auf \mathfrak{G} , so ist $d\bar{A}$ eine Zerlegung auf \mathfrak{G}^* .

Ist \bar{A} eine erzeugende Zerlegung, so ist auch $d\bar{A}$ eine erzeugende Zerlegung. Denn für $d\bar{a}, d\bar{b} \in d\bar{A}$ gibt es Elemente $\bar{a}, \bar{b}, \bar{c} \in \bar{A}$, für die $\bar{a}\bar{b} \subset \bar{c}$ gilt, und es ist $d\bar{a} \cdot d\bar{b} = d\bar{a}\bar{b} \subset d\bar{c} \in d\bar{A}$.

Ist $\overline{\mathfrak{A}}$ ein Faktoroid in \mathfrak{G} , so wird mit $d\overline{\mathfrak{A}}$ das Faktoroid in \mathfrak{G}^* bezeichnet, dessen Feld $d\bar{A}$ ist, wobei \bar{A} das Feld von $\overline{\mathfrak{A}}$ bedeutet.

Wir werden nun insbesondere isomorphe Abbildungen von Faktoroiden betrachten.

Es seien $\overline{\mathfrak{A}}, \overline{\mathfrak{C}}$ Faktoroiden in \mathfrak{G} und i ein Isomorphismus von $\overline{\mathfrak{A}}$ auf $\overline{\mathfrak{C}}$. Wir betrachten die durch ein auf $\overline{\mathfrak{A}}$ liegendes Faktoroid $\overline{\mathfrak{A}}$ erzwungene Überdeckung \mathfrak{A} von $\overline{\mathfrak{A}}$. Das auf $\overline{\mathfrak{C}}$ liegende Faktoroid $\overline{\mathfrak{C}} = i\overline{\mathfrak{A}}$ von $\overline{\mathfrak{C}}$ erzwingt eine Überdeckung \mathfrak{C} von $\overline{\mathfrak{C}}$. Für jedes Element $\dot{a} \in \mathfrak{A}$ ($\dot{c} \in \mathfrak{C}$) gibt es also ein Element $\bar{a} \in \overline{\mathfrak{A}}$, so daß $\dot{a} = s\bar{a}$ ($\dot{c} = s(i\bar{a})$) ist.

1. Die Abbildung von \mathfrak{A} auf \mathfrak{C} , in der jedes Element $\dot{a} = s\bar{a} \in \mathfrak{A}$ auf das Element $\dot{c} = s(i\bar{a}) \in \mathfrak{C}$ abgebildet wird, ist ein Isomorphismus.

Beweis. Die beschriebene Abbildung von \mathfrak{A} auf \mathfrak{C} bezeichnen wir mit \dot{i} . Die Abbildung $i_1(i_2)$ von $\overline{\mathfrak{A}}$ auf $\overline{\mathfrak{A}}$ (von $\overline{\mathfrak{C}}$ auf $\overline{\mathfrak{C}}$), in der jedes Element $\bar{a} \in \overline{\mathfrak{A}}$ ($\bar{c} \in \overline{\mathfrak{C}}$) auf das Element $\dot{a} = s\bar{a} \in \mathfrak{A}$ ($\dot{c} = s\bar{c} \in \mathfrak{C}$) abgebildet wird, ist ein Isomorphismus von $\overline{\mathfrak{A}}$ auf $\overline{\mathfrak{A}}$ (von $\overline{\mathfrak{C}}$ auf $\overline{\mathfrak{C}}$). Desgleichen ist die Abbildung i_3 von $\overline{\mathfrak{C}}$ auf $i^{-1}\overline{\mathfrak{C}} = \overline{\mathfrak{A}}$, in der jedem Elemente $\bar{c} \in \overline{\mathfrak{C}}$ das Element $i^{-1}\bar{c} = \bar{a} \in \overline{\mathfrak{A}}$ entspricht, ein Isomorphismus. Die Behauptung folgt nun aus der Gleichheit $\dot{i} = i_2 i_3^{-1} i_1^{-1}$.

2. Es sei \mathfrak{A} ein Untergruppoid in \mathfrak{G} und $\overline{\mathfrak{G}}$ ein auf \mathfrak{G} liegendes Faktoroid. Die Abbildung von $\mathfrak{A} \cap \overline{\mathfrak{G}}$ auf $\mathfrak{A} \subset \mathfrak{G}$, in der jedes Element in $\mathfrak{A} \cap \overline{\mathfrak{G}}$

auf das mit ihm inzidente Element in $\mathfrak{A} \cap \overline{\mathfrak{G}}$ abgebildet wird, ist ein Isomorphismus.

Beweis. Zu jedem Elemente $a' \in \mathfrak{A} \cap \overline{\mathfrak{G}}$ ($a'' \in \mathfrak{A} \cap \overline{\mathfrak{G}}$) gibt es genau ein mit ihm inzidentes Element $a'' \in \mathfrak{A} \cap \overline{\mathfrak{G}}$ ($a' \in \mathfrak{A} \cap \overline{\mathfrak{G}}$) und es ist $a' = A \cap a''$. Die Abbildung $a' \rightarrow a''$ von $\mathfrak{A} \cap \overline{\mathfrak{G}}$ auf $\mathfrak{A} \cap \overline{\mathfrak{G}}$ ist also umkehrbar eindeutig. Aus $a' \rightarrow a''$, $b' \rightarrow b''$ folgt $a' b' \subset a' \cdot b'$ und $a' b' \subset a'' b'' \subset a'' \cdot b''$. Also ist $a' \cdot b' = A \cap a'' \cdot b''$ und daher $a' \cdot b' \rightarrow a'' \cdot b''$.

Aus dem Satze 2 folgt leicht der Isomorphiesatz für verknüpfte Faktoroiden:

3. Es seien $\overline{\mathfrak{A}}, \overline{\mathfrak{C}}$ verknüpfte Faktoroiden in \mathfrak{G} . Die Abbildung von $\overline{\mathfrak{A}}$ auf $\overline{\mathfrak{C}}$, in der jedes Element in $\overline{\mathfrak{A}}$ auf das mit ihm inzidente Element in $\overline{\mathfrak{C}}$ abgebildet wird, ist ein Isomorphismus.

III. Ketten von Faktoroiden.

11. Grundbegriffe. Es seien $\overline{\mathfrak{A}}, \overline{\mathfrak{C}}$ Faktoroiden in \mathfrak{G} . Wir nennen das Faktoroid $\overline{\mathfrak{C}}$ eingegliedert in $\overline{\mathfrak{A}}$, wenn $\overline{\mathfrak{A}}$ das Untergruppoid $s\overline{\mathfrak{C}}$ als Element enthält, also $\overline{\mathfrak{A}} \supset s\overline{\mathfrak{C}}$; wir schreiben dann $\overline{\mathfrak{A}} \rightarrow \overline{\mathfrak{C}}$.

Es seien $\mathfrak{A}, \mathfrak{B}$ Untergruppoiden in \mathfrak{G} . Unter einer Kette von Faktoroiden von \mathfrak{A} nach \mathfrak{B} , kürzer: Kette von \mathfrak{A} nach \mathfrak{B} , verstehen wir eine geordnete endliche Menge von α (≥ 1) Faktoroiden $\overline{\mathfrak{A}}_0, \dots, \overline{\mathfrak{A}}_{\alpha-1}$ mit den folgenden Eigenschaften: 1. $\overline{\mathfrak{A}}_0$ liegt auf \mathfrak{A} ; 2. Es gilt $\overline{\mathfrak{A}}_{\beta-1} \rightarrow \overline{\mathfrak{A}}_{\beta}$ für $1 \leq \beta \leq \alpha - 1$; 3. $\overline{\mathfrak{A}}_{\alpha-1} \supset \mathfrak{B}$. Eine solche Kette wird mit

$$\overline{\mathfrak{A}}_0 \rightarrow \dots \rightarrow \overline{\mathfrak{A}}_{\alpha-1}$$

oder kürzer mit $[\overline{\mathfrak{A}}]$ bezeichnet. Für $0 \leq \beta \leq \alpha - 1$ bezeichnen wir $s\overline{\mathfrak{A}}_{\beta}$ mit \mathfrak{A}_{β} ; es ist also insbesondere $\mathfrak{A}_0 = \mathfrak{A}$. Außerdem setzen wir $\mathfrak{A}_{\alpha} = \mathfrak{B}$. Die $\mathfrak{A}, \mathfrak{B}$ sind die Enden der Kette $[\overline{\mathfrak{A}}]$. Wir sagen, die Kette $[\overline{\mathfrak{A}}]$ liege in \mathfrak{G} und im Falle $\mathfrak{A} = \mathfrak{G}$ auf \mathfrak{G} . Unter der Länge der Kette $[\overline{\mathfrak{A}}]$ verstehen wir die Anzahl α der Faktoroiden in $[\overline{\mathfrak{A}}]$. Aus der Definition der Kette $[\overline{\mathfrak{A}}]$ folgt $\mathfrak{A} = \mathfrak{A}_0 \supset \mathfrak{A}_1 \supset \dots \supset \mathfrak{A}_{\alpha} = \mathfrak{B}$.

Es sei $([\overline{\mathfrak{A}}]) = \overline{\mathfrak{A}}_0 \rightarrow \dots \rightarrow \overline{\mathfrak{A}}_{\alpha-1}$ eine Kette von \mathfrak{A} nach \mathfrak{B} von der Länge $\alpha \geq 1$. Wenn ein Faktoroid $\overline{\mathfrak{A}}_{\beta}$ ($0 \leq \beta \leq \alpha - 1$) das größte Faktoroid auf \mathfrak{A}_{β} ist, wenn also das Feld \overline{A}_{β} von $\overline{\mathfrak{A}}_{\beta}$ aus dem einzigen Elemente $A_{\beta} = A_{\beta+1}$ besteht, so wird $\overline{\mathfrak{A}}_{\beta}$ unwesentlich genannt. Andernfalls ist $\overline{\mathfrak{A}}_{\beta}$ wesentlich. $\overline{\mathfrak{A}}$ ist also wesentlich dann und nur dann, wenn $\overline{\mathfrak{A}}_{\beta+1}$ in \mathfrak{A}_{β} echt ist. Existiert in $[\overline{\mathfrak{A}}]$ wenigstens ein unwesentliches Faktoroid $\overline{\mathfrak{A}}_{\beta}$, so heißt $[\overline{\mathfrak{A}}]$ (wegen $\mathfrak{A}_{\beta} = \mathfrak{A}_{\beta+1}$) mit Wiederholungen. In diesem Falle, wenn $\alpha > 1$, kann die Kette $[\overline{\mathfrak{A}}]$ durch Streichen von $\overline{\mathfrak{A}}_{\beta}$ verkürzt werden. Sind alle Faktoroiden in der Kette $[\overline{\mathfrak{A}}]$ wesentlich, so heißt $[\overline{\mathfrak{A}}]$ ohne Wiederholungen. Die Anzahl α' wesentlicher Faktoroiden in der Kette $[\overline{\mathfrak{A}}]$ ist die

reduzierte Länge der Kette $[\bar{\mathfrak{A}}]$. Es ist $0 \leq \alpha' \leq \alpha$, wobei die Gleichheit $\alpha' = \alpha$ Ketten ohne Wiederholungen charakterisiert. Im Falle $\mathfrak{B} = \mathfrak{A}$ sind alle Faktoroiden in $[\bar{\mathfrak{A}}]$ unwesentlich, so daß $\alpha' = 0$ und umgekehrt. Wenn \mathfrak{B} in \mathfrak{A} echt ist, so kann die Kette $[\bar{\mathfrak{A}}]$ durch Streichen aller unwesentlichen Faktoroiden *reduziert*, d. h. auf eine Kette $[\bar{\mathfrak{A}}']$ ohne Wiederholungen von \mathfrak{A} nach \mathfrak{B} verkürzt werden. Die Länge der reduzierten Kette $[\bar{\mathfrak{A}}']$ ist gleich der reduzierten Länge α' der Kette $[\bar{\mathfrak{A}}]$.

Umgekehrt kann die Kette $[\bar{\mathfrak{A}}]$ *verlängert* werden, und zwar durch Eingliederung eines unwesentlichen Faktoroides zwischen zwei beliebige Faktoroiden $\bar{\mathfrak{A}}_\beta, \bar{\mathfrak{A}}_{\beta+1}$ bzw. vor (hinter) das Faktoroid $\bar{\mathfrak{A}}_0$ ($\bar{\mathfrak{A}}_{\alpha-1}$) der Kette, und zwar durch Eingliederung des größten Faktoroides auf $\bar{\mathfrak{A}}_{\beta+1}$ bzw. auf $\bar{\mathfrak{A}}_0$ ($\bar{\mathfrak{A}}_\alpha$) oder einer beliebigen endlichen Anzahl gleicher solcher Faktoroiden. Jede Verlängerung (Verkürzung) von $[\bar{\mathfrak{A}}]$ entsteht ersichtlich durch sukzessive Eingliederung (Streichen) je eines größten Faktoroides auf einem der Untergruppoiden $\bar{\mathfrak{A}}_0, \dots, \bar{\mathfrak{A}}_\alpha$. Es ist auch klar, daß jede verlängerte oder verkürzte Kette von $[\bar{\mathfrak{A}}]$ dieselbe reduzierte Länge wie die Kette $[\bar{\mathfrak{A}}]$ besitzt.

12. Elementare Ketten. Es sei $\bar{\mathfrak{A}} \supset \mathfrak{B}$ ein Faktoroid in \mathfrak{G} . Unter einer *elementaren Kette von \mathfrak{A} nach \mathfrak{B} über $\bar{\mathfrak{A}}$* , kürzer: *elementaren Kette über $\bar{\mathfrak{A}}$* , verstehen wir eine Kette von Faktoroiden von \mathfrak{A} nach \mathfrak{B}

$$([\bar{\mathfrak{A}}] \Rightarrow) \bar{\mathfrak{A}}_0 \rightarrow \dots \rightarrow \bar{\mathfrak{A}}_{\alpha-1}$$

von der Beschaffenheit, daß für $0 \leq \beta \leq \alpha - 1$ jedes Faktoroid $\bar{\mathfrak{A}}_\beta$ das Faktoroid $\bar{\mathfrak{A}}_\beta \cap \bar{\mathfrak{A}}$ überdeckt.

In einer solchen Kette ist also zuerst das Faktoroid $\bar{\mathfrak{A}}_0$ eine Überdeckung von $\bar{\mathfrak{A}}$ und es ist $\bar{\mathfrak{A}}_0 \supset \bar{\mathfrak{A}}_1 \supset \mathfrak{B}$. Nach dem Satze 8'3 ist durch diese Beziehungen das Untergruppoid $s\bar{\mathfrak{A}}_1 (\equiv \bar{\mathfrak{A}}_1)$ in $\bar{\mathfrak{A}}$ eindeutig bestimmt und das Faktoroid $(\bar{\mathfrak{A}}_1 \Rightarrow) \bar{\mathfrak{A}}_1 \cap \bar{\mathfrak{A}} \supset \mathfrak{B}$ ist ein Untergruppoid in $\bar{\mathfrak{A}}$. Weiter ist im Falle $\alpha > 1$ das Faktoroid $\bar{\mathfrak{A}}_1$ eine Überdeckung von $\bar{\mathfrak{A}}_1$ und es ist $\bar{\mathfrak{A}}_1 \supset s\bar{\mathfrak{A}}_2 \supset \mathfrak{B}$. Nach demselben Satze 8'3 ist durch diese Beziehungen das Untergruppoid $s\bar{\mathfrak{A}}_2 (\equiv \bar{\mathfrak{A}}_2)$ in $\bar{\mathfrak{A}}_1$ eindeutig bestimmt und das Faktoroid $(\bar{\mathfrak{A}}_2 \Rightarrow) \bar{\mathfrak{A}}_2 \cap \bar{\mathfrak{A}} \supset \mathfrak{B}$ ist ein Untergruppoid in $\bar{\mathfrak{A}}_1$. Weiter ist im Falle $\alpha > 2$ das Faktoroid $\bar{\mathfrak{A}}_2$ eine Überdeckung von $\bar{\mathfrak{A}}_2$ usw. Für $\alpha \geq 1$ ist das Faktoroid $\bar{\mathfrak{A}}_{\alpha-1}$ eine Überdeckung von $\bar{\mathfrak{A}}_{\alpha-1}$ und es enthält \mathfrak{B} als Element.

Zu jeder natürlichen Zahl α kann man eine elementare Kette von \mathfrak{A} nach \mathfrak{B} über $\bar{\mathfrak{A}}$ von der Länge α konstruieren in folgender Weise: Man wählt für $\bar{\mathfrak{A}}_0, \dots, \bar{\mathfrak{A}}_{\alpha-2}$ beliebige Überdeckungen von $\bar{\mathfrak{A}}, \dots, \bar{\mathfrak{A}}_{\alpha-2}$ und für $\bar{\mathfrak{A}}_{\alpha-1}$ die *kleinste* Überdeckung von $\bar{\mathfrak{A}}_{\alpha-1}$, d. h. also das Faktoroid $\bar{\mathfrak{A}}_{\alpha-1}$ selbst. Diese letzte Wahl bringt die Beziehung $\bar{\mathfrak{A}}_{\alpha-1} \supset \mathfrak{B}$ mit sich.

Es sei $0 \leq \beta \leq \alpha - 1$. Eine elementare Kette $[\bar{\mathfrak{A}}]$ von \mathfrak{A} nach \mathfrak{B} über $\bar{\mathfrak{A}}$ von der Länge α heißt *elementare Kette* (β) *von* \mathfrak{A} *nach* \mathfrak{B} *über* $\bar{\mathfrak{A}}$, kürzer: *elementare Kette* (β) *über* $\bar{\mathfrak{A}}$, wenn sie vom folgenden Typus ist:

$$\begin{aligned} \bar{\mathfrak{A}}_0 = \dots = \bar{\mathfrak{A}}_{\beta-1} &\text{ ist das größte Faktoroid } \bar{\mathfrak{A}}_{\max} \text{ von } \mathfrak{A}, \\ \bar{\mathfrak{A}}_{\beta} &\text{ ist eine Überdeckung von } \bar{\mathfrak{A}} \text{ und enthält } \mathfrak{B} \text{ als Element,} \\ \bar{\mathfrak{A}}_{\beta+1} = \dots = \bar{\mathfrak{A}}_{\alpha-1} &\text{ ist das größte Faktoroid } \bar{\mathfrak{B}}_{\max} \text{ von } \mathfrak{B}. \end{aligned}$$

In dieser Definition liest man im Falle $\beta = 0$ ($\beta = \alpha - 1$) nur die letzten (ersten) zwei Zeilen.

Wenn insbesondere $\bar{\mathfrak{A}}_{\beta} = \bar{\mathfrak{A}}$, so nennen wir $[\bar{\mathfrak{A}}]$ *ausgezeichnete elementare Kette* (β) *von* \mathfrak{A} *nach* \mathfrak{B} *über* $\bar{\mathfrak{A}}$, kürzer: *ausgezeichnete elementare Kette* (β) *über* $\bar{\mathfrak{A}}$. Insbesondere ist $[\bar{\mathfrak{A}}]$ die ausgezeichnete elementare Kette (0) von \mathfrak{A} nach \mathfrak{B} über $\bar{\mathfrak{A}}$ von der Länge 1.

Ersetzt man einerseits das Faktoroid $\bar{\mathfrak{A}}$ durch eine ausgezeichnete Kette (β) über $\bar{\mathfrak{A}}$ von der Länge α und setzt man andererseits vor $\bar{\mathfrak{A}}$ β gleiche Faktoroiden $\bar{\mathfrak{A}}_{\max}$ und hinter $\bar{\mathfrak{A}}$ $\alpha - \beta - 1$ gleiche Faktoroiden $\bar{\mathfrak{B}}_{\max}$, so erhält man dasselbe Resultat.

Es sei

$$([\bar{\mathfrak{A}}]) = \bar{\mathfrak{A}}_0 \rightarrow \dots \rightarrow \bar{\mathfrak{A}}_{\alpha-1}$$

eine elementare Kette von \mathfrak{A} nach \mathfrak{B} über $\bar{\mathfrak{A}}$.

Im Falle $\bar{\mathfrak{A}} = \bar{\mathfrak{A}}_{\max}$ gelten die Gleichheiten $\bar{\mathfrak{A}} = \bar{\mathfrak{A}}_0 = \dots = \bar{\mathfrak{A}}_{\alpha-1}$ und umgekehrt. Ist nämlich $\bar{\mathfrak{A}}$ das größte Faktoroid von \mathfrak{A} , so ist es auch jede seiner Überdeckungen, so daß im besonderen $\bar{\mathfrak{A}}_0 = \bar{\mathfrak{A}}$. Daraus folgt $\bar{\mathfrak{A}}_1 = \bar{\mathfrak{A}}$, $\bar{\mathfrak{A}}_2 = \bar{\mathfrak{A}}$ und $\bar{\mathfrak{A}}_3 = \bar{\mathfrak{A}}$ usw. und daraus die obigen Gleichheiten. Umgekehrt folgt aus diesen letzteren $\bar{\mathfrak{A}} = \bar{\mathfrak{A}}_0 = \dots = \bar{\mathfrak{A}}_{\alpha} = \mathfrak{B}$ und daraus $\bar{\mathfrak{A}} = \bar{\mathfrak{A}}_{\max}$. — Ist also $\bar{\mathfrak{A}} = \bar{\mathfrak{A}}_{\max}$, so ist $[\bar{\mathfrak{A}}]$ eine elementare Kette (β) über $\bar{\mathfrak{A}}$ und sogar eine ausgezeichnete elementare Kette (β) über $\bar{\mathfrak{A}}$, für jedes $(0 \leq) \beta (\leq \alpha - 1)$.

Die reduzierte Länge der elementaren Kette $[\bar{\mathfrak{A}}]$ ist 0 oder größer als 0, je nachdem $\bar{\mathfrak{A}}$ das größte Faktoroid von \mathfrak{A} ist oder nicht.

1. Die Ketten $[\bar{\mathfrak{A}}]$, $[\bar{\mathfrak{A}}']$ haben dann und nur dann dieselbe reduzierte Länge, wenn $[\bar{\mathfrak{A}}]$ eine elementare Kette (β) über $\bar{\mathfrak{A}}$ ist, wobei $0 \leq \beta \leq \alpha - 1$.

Beweis. a) Haben die Ketten $[\bar{\mathfrak{A}}]$, $[\bar{\mathfrak{A}}']$ dieselbe reduzierte Länge α' , so ist $\alpha' \leq 1$. Im Falle $\alpha' = 0$ ist $\bar{\mathfrak{A}} = \bar{\mathfrak{A}}_{\max}$; daraus folgt, daß $[\bar{\mathfrak{A}}]$ eine (sogar ausgezeichnete) elementare Kette (β) über $\bar{\mathfrak{A}}$ ist, wobei $0 \leq \beta \leq \alpha - 1$. Es sei also $\alpha' = 1$. Dann ist genau ein Faktoroid $\bar{\mathfrak{A}}_{\beta}$ in der elementaren Kette $[\bar{\mathfrak{A}}]$ wesentlich, wobei $0 \leq \beta \leq \alpha - 1$. Ist $\beta > 0$ ($\beta < \alpha - 1$), so sind die Faktoroiden $\bar{\mathfrak{A}}_0, \dots, \bar{\mathfrak{A}}_{\beta-1}$ ($\bar{\mathfrak{A}}_{\beta+1}, \dots, \bar{\mathfrak{A}}_{\alpha-1}$) unwesentlich, woraus

$\mathfrak{A} \equiv \mathfrak{A}_0 = \dots = \mathfrak{A}_\beta$ ($\mathfrak{A}_{\beta+1} = \dots = \mathfrak{A}_\alpha \equiv \mathfrak{B}$) folgt. Also ist $\mathfrak{A}_0 = \dots = \mathfrak{A}_{\beta-1}$ ($\mathfrak{A}_{\beta+1} = \dots = \mathfrak{A}_{\alpha-1}$) das größte Faktoroid von \mathfrak{A} (\mathfrak{B}) und \mathfrak{A}_β ist eine Überdeckung von \mathfrak{A} und enthält \mathfrak{B} als Element.

b) Ist $[\mathfrak{A}]$ eine elementare Kette (β) über \mathfrak{A} , wobei $0 \leq \beta \leq \alpha - 1$, so ist die reduzierte Länge von $[\mathfrak{A}]$ gleich 0 oder 1, je nachdem $\mathfrak{B} = \mathfrak{A}$ oder \mathfrak{B} in \mathfrak{A} echt ist. Im ersten (zweiten) Falle ist aber die reduzierte Länge von $\{\mathfrak{A}\}$ gleich 0 (1).

13. Elementare Ketten über relativ einfachen und einfachen Faktoroiden.

Es sei $\mathfrak{A} \supset \mathfrak{B}$ ein Faktoroid in \mathfrak{G} .

1. Wenn \mathfrak{A} in bezug auf \mathfrak{B} einfach ist, so ist jede elementare Kette von \mathfrak{A} nach \mathfrak{B} über \mathfrak{A} eine elementare Kette (β) über \mathfrak{A} und umgekehrt.

Beweis. a) Das Faktoroid \mathfrak{A} sei in bezug auf \mathfrak{B} einfach. Es sei

$$([\mathfrak{A}] \equiv) \mathfrak{A}_0 \rightarrow \dots \rightarrow \mathfrak{A}_{\alpha-1}$$

eine elementare Kette von \mathfrak{A} nach \mathfrak{B} über \mathfrak{A} . Ist $\mathfrak{A} = \mathfrak{A}_{\max}$, so ist $[\mathfrak{A}]$ eine (sogar ausgezeichnete) elementare Kette (β) über \mathfrak{A} für jedes ($0 \leq \beta \leq \alpha - 1$). Es sei also \mathfrak{B} in \mathfrak{A} echt. Für ein bestimmtes ($0 \leq \beta \leq \alpha - 1$) gelten dann die Gleichheiten $\mathfrak{A} \equiv \mathfrak{A}_0 = \dots = \mathfrak{A}_\beta$, während $\mathfrak{A}_{\beta+1}$ in \mathfrak{A}_β echt ist. Daraus folgt $\mathfrak{A}_0 = \dots = \mathfrak{A}_{\beta-1} = \mathfrak{A}_{\max}$, $\mathfrak{A}_\beta (\equiv \mathfrak{A}_\beta \cap \mathfrak{A}) = \mathfrak{A}$. Also ist das Faktoroid \mathfrak{A}_β eine Überdeckung von \mathfrak{A} , und nach § 2 enthält es \mathfrak{B} als Element. Ist $\beta < \alpha - 1$, so ist $\mathfrak{B} = \mathfrak{A}_{\beta+1} \supset \dots \supset \mathfrak{A}_{\alpha-1} \supset \mathfrak{B}$, und daraus folgt $\mathfrak{A}_{\beta+1} = \dots = \mathfrak{A}_{\alpha-1} = \mathfrak{B}_{\max}$.

b) Ist das Faktoroid \mathfrak{A} in bezug auf \mathfrak{B} nicht einfach, so gibt es nach § 2 eine Überdeckung \mathfrak{A}_0 von \mathfrak{A} von der Beschaffenheit, daß $\mathfrak{A} \supset \mathfrak{A}_1 \supset \mathfrak{B}$, wobei $\mathfrak{A}_1(\mathfrak{B})$ in $\mathfrak{A}(\mathfrak{A}_1)$ echt ist. Wir setzen $\mathfrak{A}_1 = \mathfrak{A}_1 (\equiv \mathfrak{A}_1 \cap \mathfrak{A})$; es ist also $\mathfrak{A}_1 \supset \mathfrak{B}$. Die Kette $\mathfrak{A}_0 \rightarrow \mathfrak{A}_1$ von \mathfrak{A} nach \mathfrak{B} über \mathfrak{A} ist elementar, aber für kein β ist sie elementar (β).

Ähnlich beweist man den folgenden Satz:

2. Ist \mathfrak{A} einfach, so ist jede elementare Kette von \mathfrak{A} nach \mathfrak{B} über \mathfrak{A} eine ausgezeichnete elementare Kette (β) über \mathfrak{A} und umgekehrt.

14. Verfeinerungen von Ketten. Es sei $\mathfrak{G} \supset \mathfrak{A} \supset \mathfrak{B}$ und sei

$$([\mathfrak{A}] \equiv) \mathfrak{A}_0 \rightarrow \dots \rightarrow \mathfrak{A}_{\alpha-1}$$

eine Kette von \mathfrak{A} nach \mathfrak{B} von der Länge $\alpha \geq 1$.

Unter einer Verfeinerung der Kette $[\mathfrak{A}]$ verstehen wir eine Kette von \mathfrak{A} nach \mathfrak{B}

$$([\mathfrak{A}] \equiv) \mathfrak{A}_{0,0} \rightarrow \dots \rightarrow \mathfrak{A}_{0,\beta_0-1} \rightarrow \mathfrak{A}_{1,0} \rightarrow \dots \rightarrow \mathfrak{A}_{1,\beta_1-1} \rightarrow \dots \rightarrow \mathfrak{A}_{\alpha-1,0} \rightarrow \dots \rightarrow \mathfrak{A}_{\alpha-1,\beta_{\alpha-1}-1}$$

von der Beschaffenheit, daß für $\gamma = 0, \dots, \alpha - 1$ die Teilkette

$$\mathfrak{A}_{\gamma, 0} \rightarrow \dots \rightarrow \mathfrak{A}_{\gamma, \beta_\gamma - 1}$$

eine elementare Kette von \mathfrak{A}_γ nach $\mathfrak{A}_{\gamma+1}$ über $\overline{\mathfrak{A}}_\gamma$ ist.

Man erhält also eine jede Verfeinerung der Kette $[\overline{\mathfrak{A}}]$, indem man jedes Faktoroid $\overline{\mathfrak{A}}_\gamma$ in der Kette durch eine elementare Kette von \mathfrak{A}_γ nach $\mathfrak{A}_{\gamma+1}$ über $\overline{\mathfrak{A}}_\gamma$ ersetzt. Wenn man insbesondere jedes Faktoroid $\overline{\mathfrak{A}}_\gamma$ in der Kette durch die ausgezeichnete elementare Kette (0) von \mathfrak{A}_γ nach $\mathfrak{A}_{\gamma+1}$ über $\overline{\mathfrak{A}}_\gamma$ von der Länge 1 ersetzt, so erhält man wieder $[\overline{\mathfrak{A}}]$. Die Kette $[\overline{\mathfrak{A}}]$ ist also zugleich ihre eigene Verfeinerung.

Die reduzierte Länge jeder Verfeinerung von $[\overline{\mathfrak{A}}]$ ist gleich der Summe reduzierter Längen der einzelnen elementaren Ketten. Dieselbe ist also gleich oder größer als die reduzierte Länge von $[\overline{\mathfrak{A}}]$. Jede Verlängerung von $[\overline{\mathfrak{A}}]$ ist eine Verfeinerung.

15. **Isomorphe Ketten.** Zwei Ketten in \mathfrak{G}

$$([\overline{\mathfrak{A}}] \equiv) \overline{\mathfrak{A}}_0 \rightarrow \dots \rightarrow \overline{\mathfrak{A}}_{\alpha-1},$$

$$([\overline{\mathfrak{C}}] \equiv) \overline{\mathfrak{C}}_0 \rightarrow \dots \rightarrow \overline{\mathfrak{C}}_{\beta-1}$$

heißen *isomorph*, wenn es eine umkehrbar eindeutige Zuordnung — die sogenannte *erste Abbildung* — der Faktoroide in der einen Kette zu den Faktoroide in der anderen gibt von der folgenden Beschaffenheit: Je zwei einander zugeordnete Faktoroide $\overline{\mathfrak{A}}_\gamma, \overline{\mathfrak{C}}_\delta$ sind isomorph, wobei sich in dem Isomorphismus $A_{\gamma+1}$ und $C_{\delta+1}$ entsprechen. Sind $[\overline{\mathfrak{A}}], [\overline{\mathfrak{C}}]$ isomorph, so sagen wir auch $[\overline{\mathfrak{A}}] ([\overline{\mathfrak{C}}])$ sei isomorph mit $[\overline{\mathfrak{C}}] ([\overline{\mathfrak{A}}])$.

Ersichtlich haben zwei isomorphe Ketten in \mathfrak{G} dieselbe Länge.

Es seien $[\overline{\mathfrak{A}}], [\overline{\mathfrak{C}}]$ isomorphe Ketten in \mathfrak{G} von der Länge $\alpha \geq 1$. Jedem unwesentlichen Faktoroid in $[\overline{\mathfrak{A}}]$ entspricht in der ersten Abbildung wieder ein unwesentliches Faktoroid in $[\overline{\mathfrak{C}}]$. Ist $\alpha > 1$, so sind beide durch das Streichen solcher zugeordneten unwesentlichen Faktoroide verkürzte Ketten wieder isomorph. Daraus folgt, daß $[\overline{\mathfrak{A}}], [\overline{\mathfrak{C}}]$ dieselbe reduzierte Länge besitzen und die beiden reduzierten Ketten wieder isomorph sind.

Es seien $\mathfrak{A} \ni \mathfrak{B}, \mathfrak{C} \ni \mathfrak{D}$ Faktoroide in \mathfrak{G} . Wir setzen voraus, daß ein Isomorphismus i von \mathfrak{A} auf \mathfrak{C} existiert, wobei $i\mathfrak{B} = \mathfrak{D}$. Die beiden Ketten $\{\overline{\mathfrak{A}}\}, \{\overline{\mathfrak{C}}\}$ sind also isomorph. Wir betrachten weiter eine elementare Kette von \mathfrak{A} nach \mathfrak{B} über $\overline{\mathfrak{A}}$:

$$([\overline{\mathfrak{A}}] \equiv) \mathfrak{A}_0 \rightarrow \dots \rightarrow \mathfrak{A}_{\alpha-1}.$$

1. *Es gibt eine elementare Kette $[\overline{\mathfrak{C}}]$ von \mathfrak{C} nach \mathfrak{D} über $\overline{\mathfrak{C}}$, die mit $[\overline{\mathfrak{A}}]$ isomorph ist.* Dieselbe ist durch die im Teile a) des folgenden Beweises beschriebene Konstruktion bestimmt.

Beweis. a) Nach der Definition der elementaren Kette ist das Faktoroid $\bar{\mathfrak{A}}_\gamma$ für $\gamma = 0, \dots, \alpha - 1$, eine Überdeckung von $\mathfrak{A}_\gamma \cap \bar{\mathfrak{A}} (\equiv \bar{\mathfrak{A}}_\gamma)$. Es sei $\bar{\mathfrak{A}}_\gamma$ das die Überdeckung $\bar{\mathfrak{A}}_\gamma$ von \mathfrak{A}_γ erzwingende Faktoroid von $\bar{\mathfrak{A}}_\gamma$. Es ist also insbesondere $\bar{\mathfrak{A}}_{\gamma+1} \in \bar{\mathfrak{A}}_\gamma$, wobei $\bar{\mathfrak{A}}_\alpha = \{\mathfrak{B}\}$ gesetzt wird. Wir definieren nun $\bar{\mathfrak{C}}_\gamma, \bar{\mathfrak{C}}_\gamma$, für $\gamma = 0, \dots, \alpha - 1$ in folgender Weise: $\bar{\mathfrak{C}}_\gamma = i\bar{\mathfrak{A}}_\gamma$; setzen wir $\bar{\mathfrak{C}}_\gamma = s\bar{\mathfrak{C}}_\gamma$, so ist $\bar{\mathfrak{C}}_\gamma = \bar{\mathfrak{C}}_\gamma \cap \bar{\mathfrak{C}}, \bar{\mathfrak{C}}_0 = \bar{\mathfrak{C}}$. $\bar{\mathfrak{C}}_\gamma = i\bar{\mathfrak{A}}_\gamma$; es ist also insbesondere $\bar{\mathfrak{C}}_{\gamma+1} (= i\bar{\mathfrak{A}}_{\gamma+1}) \in \bar{\mathfrak{C}}_\gamma$, wobei $\bar{\mathfrak{C}}_\alpha = \{\mathfrak{D}\}$ gesetzt wird. $\bar{\mathfrak{C}}_\gamma$ ist die durch $\bar{\mathfrak{C}}_\gamma$ erzwungene Überdeckung von $\bar{\mathfrak{C}}_\gamma$; da $\bar{\mathfrak{C}}_{\gamma+1} \in \bar{\mathfrak{C}}_\gamma$ ist, so ist $\bar{\mathfrak{C}}_{\gamma+1} \in \bar{\mathfrak{C}}_\gamma$, wobei $\bar{\mathfrak{C}}_\alpha = \mathfrak{D}$ gesetzt wird. — Es ist also

$$(\bar{\mathfrak{C}}) \equiv \bar{\mathfrak{C}}_0 \rightarrow \dots \rightarrow \bar{\mathfrak{C}}_{\alpha-1}$$

eine Kette von $\bar{\mathfrak{C}}$ nach \mathfrak{D} , und zwar eine elementare Kette über $\bar{\mathfrak{C}}$.

b) Wir definieren nun die erste Abbildung der Kette $[\bar{\mathfrak{A}}]$ auf $[\bar{\mathfrak{C}}]$, indem wir für $\gamma = 0, \dots, \alpha - 1$ dem Faktoroid $\bar{\mathfrak{A}}_\gamma$ das Faktoroid $\bar{\mathfrak{C}}_\gamma$ zuordnen. Nach dem Satze 10.1 gibt es eine isomorphe Abbildung i_γ von $\bar{\mathfrak{A}}_\gamma$ auf $\bar{\mathfrak{C}}_\gamma$, in der jedes Element $s\bar{a}$, für $\bar{a} \in \bar{\mathfrak{A}}_\gamma$ auf das Element $s(i\bar{a})$ abgebildet wird, für die also $s(i\bar{a}) = i_\gamma s\bar{a}$ und im besonderen $\bar{\mathfrak{C}}_{\gamma+1} = i_\gamma \bar{\mathfrak{A}}_{\gamma+1}$ ist. Also sind die Ketten $[\bar{\mathfrak{A}}], [\bar{\mathfrak{C}}]$ isomorph.

Ist $[\bar{\mathfrak{A}}]$ eine elementare Kette (β) (eine ausgezeichnete elementare Kette (β)) von \mathfrak{A} nach \mathfrak{B} über $\bar{\mathfrak{A}}$, so ist auch $[\bar{\mathfrak{C}}]$ eine elementare Kette (β) (eine ausgezeichnete elementare Kette (β)) von $\bar{\mathfrak{C}}$ nach \mathfrak{D} über $\bar{\mathfrak{C}}$. Denn ist $[\bar{\mathfrak{A}}]$ eine elementare Kette (β) von \mathfrak{A} nach \mathfrak{B} über $\bar{\mathfrak{A}}$ und ist $\beta > 0$ ($\beta < \alpha - 1$), so ist $\bar{\mathfrak{A}}_0 = \dots = \bar{\mathfrak{A}}_{\beta-1}$ ($\bar{\mathfrak{A}}_{\beta+1} = \dots = \bar{\mathfrak{A}}_{\alpha-1}$) das größte Faktoroid von \mathfrak{A} (\mathfrak{B}) und infolgedessen $\bar{\mathfrak{C}}_0 = \dots = \bar{\mathfrak{C}}_{\beta-1}$ ($\bar{\mathfrak{C}}_{\beta+1} = \dots = \bar{\mathfrak{C}}_{\alpha-1}$) das größte Faktoroid von $\bar{\mathfrak{C}}$ (\mathfrak{D}). Weiter ist $\bar{\mathfrak{A}}_\beta = \bar{\mathfrak{A}}, \{B\} \in \bar{\mathfrak{A}}_\beta$, woraus $\bar{\mathfrak{C}}_\beta = i\bar{\mathfrak{A}}_\beta = i\bar{\mathfrak{A}} = \bar{\mathfrak{C}}, \{D\} = \{iB\} \in i\bar{\mathfrak{A}}_\beta = \bar{\mathfrak{C}}_\beta$ folgt. $\bar{\mathfrak{C}}_\beta$ ist also eine Überdeckung von $\bar{\mathfrak{C}}$ und enthält \mathfrak{D} als Element. Ist sogar $\bar{\mathfrak{A}}_\beta = \bar{\mathfrak{A}}$, so ist $\bar{\mathfrak{A}}_\beta$ das kleinste Faktoroid von $\bar{\mathfrak{A}}$, also $\bar{\mathfrak{C}}_\beta$ das kleinste Faktoroid von $\bar{\mathfrak{C}}$, woraus $\bar{\mathfrak{C}}_\beta = \bar{\mathfrak{C}}$ folgt.

Es seien $[\bar{\mathfrak{A}}], [\bar{\mathfrak{C}}]$ isomorphe Ketten in \mathfrak{G} von der Länge $\alpha \geq 1$.

2. Zu jeder Verfeinerung von $[\bar{\mathfrak{A}}]$ gibt es eine isomorphe Verfeinerung von $[\bar{\mathfrak{C}}]$.

Beweis. Es sei $[\bar{\mathfrak{A}}]$ eine Verfeinerung von $[\bar{\mathfrak{A}}]$ und für $\gamma = 0, \dots, \alpha - 1$ sei $[\bar{\mathfrak{A}}_\gamma]$ die elementare Kette von $\bar{\mathfrak{A}}_\gamma$ nach $\bar{\mathfrak{A}}_{\gamma+1}$ über $[\bar{\mathfrak{A}}_\gamma]$, die sich in der Kette $[\bar{\mathfrak{A}}]$ befindet. Da nach der Voraussetzung die Kette $[\bar{\mathfrak{A}}]$ mit $[\bar{\mathfrak{C}}]$ isomorph ist, so entspricht in der ersten Abbildung jedem Faktoroid $\bar{\mathfrak{A}}_\gamma \in [\bar{\mathfrak{A}}]$ umkehrbar eindeutig ein Faktoroid $\bar{\mathfrak{C}}_\beta \in [\bar{\mathfrak{C}}]$ und es gibt einen Isomorphismus i_γ von $\bar{\mathfrak{A}}_\gamma$ auf $\bar{\mathfrak{C}}_\beta$, für den $i_\gamma \bar{\mathfrak{A}}_{\gamma+1} = \bar{\mathfrak{C}}_{\beta+1}$ ist. Es gibt also nach dem Satze 1 eine mit $[\bar{\mathfrak{A}}_\gamma]$ isomorphe elementare Kette $[\bar{\mathfrak{C}}_\beta]$ von $\bar{\mathfrak{C}}_\beta$

nach $\mathbb{C}_{\beta+1}$ über $\bar{\mathbb{C}}_{\beta}$. Ersetzt man in der Kette $[\bar{\mathbb{C}}]$ für $\gamma = 0, \dots, \alpha - 1$, das Faktoroid $\bar{\mathbb{C}}_{\beta}$ durch die elementare Kette $[\bar{\mathbb{C}}_{\beta}]$, so erhält man eine mit $[\bar{\mathbb{A}}]$ isomorphe Verfeinerung von $[\bar{\mathbb{C}}]$.

16. Relativ einfache und einfache Ketten. Es sei $\mathbb{A} \supset \mathbb{B}$ und

$$([\bar{\mathbb{A}}] \Rightarrow) \bar{\mathbb{A}}_0 \rightarrow \dots \rightarrow \bar{\mathbb{A}}_{\alpha-1}$$

eine Kette von \mathbb{A} nach \mathbb{B} . Wir nennen die Kette $[\bar{\mathbb{A}}]$ *relativ einfach*, wenn jedes Faktoroid $\bar{\mathbb{A}}_{\gamma} \in [\bar{\mathbb{A}}]$, $\gamma = 0, \dots, \alpha - 1$, in bezug auf $\mathbb{A}_{\gamma+1}$ ($\mathbb{A}_{\alpha} = \mathbb{B}$) einfach ist. Wenn jedes Faktoroid $\bar{\mathbb{A}}_{\gamma} \in [\bar{\mathbb{A}}]$ sogar einfach ist, so wird die Kette $[\bar{\mathbb{A}}]$ *einfach* genannt.

1. Dann und nur dann, wenn die Kette $[\bar{\mathbb{A}}]$ relativ einfach ist, haben alle ihre Verfeinerungen dieselbe reduzierte Länge.

Beweis. a) Die Kette $[\bar{\mathbb{A}}]$ sei relativ einfach. Wir haben zu zeigen, daß jede Verfeinerung $[\bar{\mathbb{A}}']$ von $[\bar{\mathbb{A}}]$ dieselbe reduzierte Länge wie $[\bar{\mathbb{A}}]$ besitzt. Nun erhält man aber jede Verfeinerung $[\bar{\mathbb{A}}']$ von $[\bar{\mathbb{A}}]$, indem man jedes Faktoroid $\bar{\mathbb{A}}_{\gamma}$ in der Kette $[\bar{\mathbb{A}}]$ durch eine geeignete elementare Kette $[\bar{\mathbb{A}}'_{\gamma}]$ von \mathbb{A}_{γ} nach $\mathbb{A}_{\gamma+1}$ über $\bar{\mathbb{A}}_{\gamma}$ ersetzt. Nach dem Satze 13 '1 ist jede elementare Kette $[\bar{\mathbb{A}}'_{\gamma}]$ eine elementare Kette (β) über $\bar{\mathbb{A}}_{\gamma}$ und nach 12 '1 ist ihre reduzierte Länge gleich der reduzierten Länge von $\{\bar{\mathbb{A}}_{\gamma}\}$. Außerdem ist die reduzierte Länge von $[\bar{\mathbb{A}}']$ gleich der Summe der reduzierten Längen der einzelnen Ketten $[\bar{\mathbb{A}}'_{\gamma}]$, und desgleichen ist die reduzierte Länge von $[\bar{\mathbb{A}}]$ gleich der Summe der einzelnen Ketten $\{\bar{\mathbb{A}}_{\gamma}\}$.

b) Wir setzen voraus, daß jede Verfeinerung $[\bar{\mathbb{A}}']$ von $[\bar{\mathbb{A}}]$ dieselbe reduzierte Länge wie die Kette $[\bar{\mathbb{A}}]$ besitzt. Dann ist die reduzierte Länge der elementaren Teilkette $[\bar{\mathbb{A}}'_{\gamma}] \subset [\bar{\mathbb{A}}]$ über jedem Faktoroid $\bar{\mathbb{A}}_{\gamma} \in [\bar{\mathbb{A}}]$ gleich der reduzierten Länge von $\{\bar{\mathbb{A}}_{\gamma}\}$, $\gamma = 0, \dots, \alpha - 1$. Also ist nach dem Satze 12 '1 $[\bar{\mathbb{A}}'_{\gamma}]$ eine elementare Kette (β) über $\bar{\mathbb{A}}_{\gamma}$, und nach 13 '1 ist das Faktoroid $\bar{\mathbb{A}}_{\gamma}$ einfach in bezug auf $\mathbb{A}_{\gamma+1}$.

Ähnlich ist der Beweis des folgenden Satzes:

2. Dann und nur dann, wenn die Kette $[\bar{\mathbb{A}}]$ einfach ist, ist jede Verfeinerung von $[\bar{\mathbb{A}}]$ eine Verlängerung.

Es seien $[\bar{\mathbb{A}}]$, $[\bar{\mathbb{C}}]$ isomorphe Ketten in \mathbb{G} .

3. Wenn $[\bar{\mathbb{A}}]$ relativ einfach ist, so ist es auch $[\bar{\mathbb{C}}]$.

Beweis. Wegen des Isomorphismus haben die Ketten $[\bar{\mathbb{A}}]$, $[\bar{\mathbb{C}}]$ dieselbe reduzierte Länge α' . Ist die Kette $[\bar{\mathbb{C}}]$ nicht relativ einfach, so gibt es nach 1 eine Verfeinerung $[\bar{\mathbb{C}}']$ von $[\bar{\mathbb{C}}]$, deren reduzierte Länge β' größer als α' ist. Nach 15 '2 gibt es eine mit $[\bar{\mathbb{C}}']$ isomorphe Verfeinerung $[\bar{\mathbb{A}}']$ von $[\bar{\mathbb{A}}]$. Die reduzierte Länge von $[\bar{\mathbb{A}}']$ ist nach 16 '1 gleich α' , und wegen des Isomorphismus der Ketten $[\bar{\mathbb{A}}]$, $[\bar{\mathbb{C}}]$ ist sie gleich β' .

17. Adjungierte Ketten. Es seien

$$([\bar{\mathfrak{A}}] \equiv) \bar{\mathfrak{A}}_0 \rightarrow \dots \rightarrow \bar{\mathfrak{A}}_{\alpha-1},$$

$$([\bar{\mathfrak{C}}] \equiv) \bar{\mathfrak{C}}_0 \rightarrow \dots \rightarrow \bar{\mathfrak{C}}_{\beta-1},$$

Ketten in \mathfrak{G} . Wir sagen, daß $[\bar{\mathfrak{A}}]$, $[\bar{\mathfrak{C}}]$ *adjungierte Ketten* sind oder daß die Kette $[\bar{\mathfrak{A}}]$ ($[\bar{\mathfrak{C}}]$) zu $[\bar{\mathfrak{C}}]$ ($[\bar{\mathfrak{A}}]$) adjungiert ist, wenn 1. die Enden von $[\bar{\mathfrak{A}}]$ und $[\bar{\mathfrak{C}}]$ dieselben sind; 2. jedes Faktoroid $\bar{\mathfrak{A}}_\gamma \in [\bar{\mathfrak{A}}]$ zu jedem Faktoroid $\bar{\mathfrak{C}}_\delta \in [\bar{\mathfrak{C}}]$ in bezug auf $\mathfrak{A}_{\gamma+1}$, $\mathfrak{C}_{\delta+1}$ adjungiert ist ($\gamma = 0, \dots, \alpha-1$; $\delta = 0, \dots, \beta-1$). Die Bedingung 1. drückt die folgenden Gleichheiten aus: $\mathfrak{A}_0 = \mathfrak{C}_0 (\equiv \mathfrak{A})$, $\mathfrak{A}_\alpha = \mathfrak{C}_\beta (\equiv \mathfrak{B})$. Nach 2. sind die Ketten $[\bar{\mathfrak{A}}]$, $[\bar{\mathfrak{C}}]$ nur dann adjungiert, wenn der Durchschnitt $\mathfrak{A}_\mu \cap \mathfrak{C}_\nu$ für $\mu = 0, \dots, \alpha$; $\nu = 0, \dots, \beta$ existiert.

Wenn man von zwei adjungierten Ketten eine der beiden Ketten oder beide verkürzt oder verlängert, so bekommt man wieder zwei adjungierte Ketten.

Es seien $[\bar{\mathfrak{A}}]$, $[\bar{\mathfrak{C}}]$ adjungierte Ketten in \mathfrak{G} .

1. (Verallgemeinerung des Satzes von Jordan-Hölder-Schreier.) Die Ketten $[\bar{\mathfrak{A}}]$, $[\bar{\mathfrak{C}}]$ besitzen isomorphe Verfeinerungen. Dieselben sind durch die im Teile a) des folgenden Beweises gegebene Konstruktion bestimmt.

Beweis. a) Da nach der Voraussetzung die Faktoroiden $\bar{\mathfrak{A}}_\gamma$, $\bar{\mathfrak{C}}_\delta$ für $\gamma = 0, \dots, \alpha-1$; $\delta = 0, \dots, \beta-1$ adjungiert sind, so sind

$$\bar{\mathfrak{A}}_{\gamma,\nu} \equiv \mathfrak{C}_\nu \subset \bar{\mathfrak{A}}_\gamma, \quad \bar{\mathfrak{C}}_{\delta,\mu} \equiv \mathfrak{A}_\mu \subset \bar{\mathfrak{C}}_\delta,$$

für $\mu = 0, \dots, \alpha$; $\nu = 0, \dots, \beta$, Faktoroiden in \mathfrak{G} . Wir setzen $\bar{\mathfrak{A}}_{\gamma,\nu} = s \bar{\mathfrak{A}}_{\gamma,\nu}$, $\bar{\mathfrak{C}}_{\delta,\mu} = s \bar{\mathfrak{C}}_{\delta,\mu}$. Da die Ketten $[\bar{\mathfrak{A}}]$, $[\bar{\mathfrak{C}}]$ dieselben Enden haben, so ist

$$\begin{aligned} \bar{\mathfrak{A}}_{\gamma,0} &= \bar{\mathfrak{A}}_\gamma, & \bar{\mathfrak{A}}_{\gamma,\beta} &= \bar{\mathfrak{A}}_{\gamma+1}, \\ \bar{\mathfrak{C}}_{\delta,0} &= \bar{\mathfrak{C}}_\delta, & \bar{\mathfrak{C}}_{\delta,\alpha} &= \bar{\mathfrak{C}}_{\delta+1}. \end{aligned}$$

Nach 8.1 gibt es verknüpfte Überdeckungen $\bar{\mathfrak{A}}_{\gamma,\delta}$, $\bar{\mathfrak{C}}_{\delta,\gamma}$ von $\bar{\mathfrak{A}}_{\gamma,\delta}$, $\bar{\mathfrak{C}}_{\delta,\gamma}$, so daß $\bar{\mathfrak{A}}_{\gamma,\delta+1} \in \bar{\mathfrak{A}}_{\gamma,\delta}$, $\bar{\mathfrak{C}}_{\delta,\gamma+1} \in \bar{\mathfrak{C}}_{\delta,\gamma}$. Die Untergruppoiden $\bar{\mathfrak{A}}_{\gamma,\delta+1}$, $\bar{\mathfrak{C}}_{\delta,\gamma+1}$ sind inzident. Da für $0 \leq \delta \leq \beta-1$ das Faktoroid $\bar{\mathfrak{A}}_{\gamma,\delta}$ eine Überdeckung von $\bar{\mathfrak{A}}_{\gamma,\delta} (= \bar{\mathfrak{A}}_{\gamma,\delta} \cap \bar{\mathfrak{A}}_\gamma)$ ist, so ist

$$\bar{\mathfrak{A}}_{\gamma,0} \rightarrow \dots \rightarrow \bar{\mathfrak{A}}_{\gamma,\beta-1}$$

eine elementare Kette von $\bar{\mathfrak{A}}_\gamma$ nach $\bar{\mathfrak{A}}_{\gamma+1}$ über $\bar{\mathfrak{A}}_\gamma$ und desgleichen ist

$$\bar{\mathfrak{C}}_{\delta,0} \rightarrow \dots \rightarrow \bar{\mathfrak{C}}_{\delta,\alpha-1}$$

eine elementare Kette von $\bar{\mathfrak{C}}_\delta$ nach $\bar{\mathfrak{C}}_{\delta+1}$ über $\bar{\mathfrak{C}}_\delta$. Die Ketten von \mathfrak{A} nach \mathfrak{B}

$$([\bar{\mathfrak{A}}] \equiv) \bar{\mathfrak{A}}_{0,0} \rightarrow \dots \rightarrow \bar{\mathfrak{A}}_{0,\beta-1} \rightarrow \bar{\mathfrak{A}}_{1,0} \rightarrow \dots \rightarrow \bar{\mathfrak{A}}_{\alpha-1,\beta-1},$$

$$([\bar{\mathfrak{C}}] \equiv) \bar{\mathfrak{C}}_{0,0} \rightarrow \dots \rightarrow \bar{\mathfrak{C}}_{0,\alpha-1} \rightarrow \bar{\mathfrak{C}}_{1,0} \rightarrow \dots \rightarrow \bar{\mathfrak{C}}_{\beta-1,\alpha-1}$$

sind die erwähnten Verfeinerungen.

b) Um zu beweisen, daß die Ketten $[\mathfrak{A}], [\mathfrak{C}]$ isomorph sind, definieren wir zuerst die erste Abbildung von $[\mathfrak{A}]$ auf $[\mathfrak{C}]$ in der Weise, daß wir jedem Faktoroid $\mathfrak{A}_{\gamma, \delta}$ das Faktoroid $\mathfrak{C}_{\delta, \gamma}$ zuordnen ($\gamma = 0, \dots, \alpha - 1$; $\delta = 0, \dots, \beta - 1$). Da die Faktoroiden $\mathfrak{A}_{\gamma, \delta}, \mathfrak{C}_{\delta, \gamma}$ verknüpft sind, so ist nach 10.3 die Abbildung $i_{\gamma, \delta}$ von $\mathfrak{A}_{\gamma, \delta}$ auf $\mathfrak{C}_{\delta, \gamma}$, in der jedes Element in $\mathfrak{A}_{\gamma, \delta}$ auf das mit ihm inzidente Element in $\mathfrak{C}_{\delta, \gamma}$ abgebildet wird, ein Isomorphismus. Weil insbesondere die $\mathfrak{A}_{\gamma, \delta+1}, \mathfrak{C}_{\delta, \gamma+1}$ inzident sind, so ist $i_{\gamma, \delta} \mathfrak{A}_{\gamma, \delta+1} = \mathfrak{C}_{\delta, \gamma+1}$.

2. Wenn eine der Ketten $[\mathfrak{A}], [\mathfrak{C}]$ relativ einfach ist, so ist ihre reduzierte Länge nicht kleiner als die reduzierte Länge der anderen.

Beweis. Wir setzen voraus, daß die Kette $[\mathfrak{A}]$ relativ einfach ist. Es sei α' (β') die reduzierte Länge von $[\mathfrak{A}]$ ($[\mathfrak{C}]$). Nach 1 gibt es isomorphe Verfeinerungen der Ketten $[\mathfrak{A}], [\mathfrak{C}]$. Dieselben haben die gleiche reduzierte Länge γ' . Nun ist aber nach 16.1 $\alpha' = \gamma'$ und außerdem ist $\beta' \leq \gamma'$. Also ist $\alpha' \geq \beta'$.

Aus diesem Satze folgt der folgende Satz:

3. Wenn die Ketten $[\mathfrak{A}], [\mathfrak{C}]$ relativ einfach sind, so sind ihre reduzierten Längen einander gleich.

(Eingegangen am 6. 9. 1940.)

Modulformen und quadratische Formen über dem quadratischen Zahlkörper $R(\sqrt{5})$.

Von

Hans Maaß in Heidelberg.

Es hat sich gezeigt¹⁾, daß die Peterssorsche Methode zur Konstruktion automorpher Formen²⁾, welche nur die Kenntnis von Multiplikatorsystemen voraussetzt, auch auf Gruppen simultaner linear gebrochener Substitutionen vom Typus der Hilbertschen Modulgruppe mit Erfolg angewendet werden kann. Der erste Schritt zur expliziten Aufstellung automorpher Formen der Dimension $-r$ zu einer vorgelegten Gruppe G wird demnach darin bestehen, daß man sich eine Übersicht über alle Multiplikatorsysteme zur Gruppe G und zur Dimension $-r$ verschafft. Bei einem derartigen Versuch stößt man aber im allgemeinen, wenn es sich um Gruppen zu mehreren Veränderlichen handelt, auf erhebliche Schwierigkeiten; denn in den seltensten Fällen sind Erzeugende geschweige denn definierende Relationen der vorgelegten Gruppe bekannt. Es erscheint fürs erste geboten, an speziellen einfachen Gruppen als Beispielen die Verhältnisse zu studieren. Ich habe zu diesem Zweck die Hilbertsche Modulgruppe M zum quadratischen Zahlkörper $R(\sqrt{5})$ ausgewählt. Als Definitionsbereich für Modulformen zu M oder zu Untergruppen U von M legen wir den Bereich

$$(1) \quad \Im \tau > 0, \Im \tau' < 0$$

zugrunde. Ein Zahlssystem $v(S)$ ($S \subset U$) heißt ein Multiplikatorsystem zu U und zur Dimension $-r$, wenn

$$(2) \quad v(S_1 S_2) = \sigma^{(r)}(S_1, S_2) v(S_1) v(S_2) \text{ für } S_1, S_2 \subset U, v(-E) = 1, \\ \text{falls } -E \subset U.$$

Dabei ist³⁾

$$(3) \quad \sigma^{(r)}(S_1, S_2) = e^{2\pi i r S(w(S_1, S_2) \operatorname{sgn} \sqrt{5})}$$

¹⁾ H. Maaß, Zur Theorie der automorphen Funktionen von n Veränderlichen. Math. Annalen 117 (1940), S. 538–578.

²⁾ H. Petersson, Theorie der automorphen Formen beliebiger reeller Dimension und ihre Darstellung durch eine neue Art Poincaréscher Reihen. Math. Annalen 103 (1930), S. 369–436.

³⁾ H. Maaß, Konstruktion ganzer Modulformen halbzahlgiger Dimension mit ϕ -Multiplikatoren in zwei Variablen. Math. Zeitschr. 43 (1938), S. 709–738.

das Faktorsystem zur Dimension $-r$ und E die Einheitstransformation. Außer M wird noch die Thetagruppe T aus M betrachtet, welche aus allen Substitutionen

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in M$$

besteht, für welche entweder

$$(4) \quad \beta \equiv \gamma \equiv 0 \pmod{2} \quad \text{oder} \quad \alpha \equiv \delta \equiv 0 \pmod{2}.$$

Im ersten Teil dieser Arbeit wird nun folgendes bewiesen. Für M gibt es nur zu ganzzahligen Dimensionen ein Multiplikatorsystem, und zwar genau eins, nämlich das triviale System $v(S) = 1$ für $S \in M$. Es liegen hier also ganz andere Verhältnisse vor als bei der rationalen Modulgruppe M_0 , bei welcher jede reelle Zahl als Dimension vorkommt²⁾. Allgemein kann leicht bewiesen werden, daß es zu jeder Hilbertschen Modulgruppe ($\neq M_0$) nur Multiplikatorsysteme zu rationalzahligen Dimensionen gibt. Es kommt hierbei die Tatsache zum Ausdruck, daß es auch hyperbolische Substitutionen mit vorgegebener parabolischer Spitze als Fixpunkt in den betrachteten Gruppen gibt⁴⁾. Bei der Thetagruppe T treten nur ganz- und halbzahlige Werte von $-r$ als Dimensionen auf. Zu jeder solchen Dimension gibt es genau vier Multiplikatorsysteme; diese können explizit angegeben werden.

An diese Betrachtung schließt sich die Berechnung von Identitäten zwischen Eisensteinreihen und Thetareihen zur Gruppe M . Für die Maximalzahl $a(r)$ linear unabhängiger Modulformen zu M von der Dimension $-r < 0$ ist die Abschätzung

$$(5) \quad a(r) < cr^2$$

mit konstantem c bewiesen worden¹⁾. Der Beweis dieses Satzes soll hier in seinen wesentlichen Teilen reproduziert werden und liefert mit einer von Herrn Witt⁵⁾ bemerkten Vereinfachung des ursprünglichen Siegelschen Beweisganges für die niederen Dimensionen die Abschätzungen

$$(6) \quad a(1) = 0, \quad a(2) = 1, \quad a(3) \leq 2, \quad a(4) \leq 3.$$

Außerdem ist dann die Möglichkeit gegeben, numerisch zu entscheiden, wann zwei Modulformen der Dimension $-r$ ($r \leq 4$) identisch sind. Auf diesem

¹⁾ H. Maaß, Über Gruppen von hyperabelschen Transformationen. Sitzungsber. der Heidelberger Akademie der Wissenschaften, math.-naturwiss. Klasse (1940), 2. Abhandlung.

⁵⁾ E. Witt, Eine Identität zwischen Modulformen zweiten Grades. Abhandl. aus d. Math. Seminar d. Hanischen Universität (im Druck).

Wege ergeben sich die nachfolgenden Identitäten. Die Eisensteinreihe $G_r(\tau)$ von \mathbf{M} zur Dimension $-r^6$ ($r \equiv 0(2)$, $r > 0$) sei so normiert, daß in ihrer Fourierentwicklung zur parabolischen Spitze ∞ das konstante Glied 1 lautet. Dann ist

$$(7) \quad (G_2(\tau))^2 = G_4(\tau).$$

Durch die Thetareihen wird zwischen der Theorie der Modulformen zu \mathbf{M} und den positiv definiten ganzzahligen geraden quadratischen Formen $Q(x_1, \dots, x_m)$ mit Determinante ε^{2r} der folgende Zusammenhang gestiftet. Wir bezeichnen mit Ω_m die Klasse der mit $Q(x_1, \dots, x_m)$ ganzzahlig äquivalenten quadratischen Formen mit der Determinante ε^{2r} ; dann erweist sich

$$(8) \quad \vartheta(\tau, \Omega_m) = \sum_{x_1, \dots, x_m} e^{\pi i S \frac{Q(x_1, \dots, x_m) \tau}{\sqrt{\varepsilon}}},$$

wobei über alle ganzen Zahlssysteme x_1, \dots, x_m aus $R(\sqrt{5})$ summiert wird, als eine Modulform zu \mathbf{M} von der Dimension $-\frac{m}{2}$. Die Zahl m der Veränderlichen muß notwendig durch 4 teilbar sein. Diese Bedingung ist auch hinreichend für die Existenz von quadratischen Formen der genannten Art; denn es wird eine nicht leere Klasse Ω_4 nachgewiesen. Für diese muß dann

$$(9) \quad \vartheta(\tau, \Omega_4) = G_2(\tau)$$

gelten. Ω_8 sei die Klasse, in welcher die bekannte quadratische Form von 8 Veränderlichen mit ganz rationalen Koeffizienten liegt; dann wird

$$(10) \quad \vartheta(\tau, \Omega_8) = G_4(\tau)$$

bewiesen, womit sich nach (7) und (9) ergibt, daß jede ganze gerade Zahl aus $R(\sqrt{5})$ durch Ω_8 und $\Omega_4 + \Omega_4$ (direkte Summe) gleich oft dargestellt wird, obgleich die beiden Klassen voneinander verschieden sind. In einer abschließenden Betrachtung wird gezeigt, daß andere Klassen als Ω_4 , $\Omega_4 + \Omega_4$, Ω_8 zur Veränderlichenzahl 4 und 8 nicht vorkommen. Der Beweis erfolgt durch Diskussion der den Klassen quadratischer Formen zugeordneten Gitter⁵⁾.

⁵⁾ E. Hecke, Analytische Funktionen und algebraische Zahlen, II. Teil. Abhandl. aus d. Math. Seminar d. Hannischen Universität 3 (1924), S. 213–236.

§ 1.

Multiplikatorsysteme zu M und T .

Wir setzen in fester Bezeichnung

$$\varepsilon = \frac{1 + \sqrt{5}}{2}.$$

Die Zahlen $1, \varepsilon$ bilden eine Basis für den Bereich \mathfrak{o} aller ganzen Zahlen aus $R(\sqrt{5})$. Bekanntlich⁷⁾ wird M von den Substitutionen

$$(11) \quad S_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad S_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad S_3 = \begin{pmatrix} \varepsilon & 0 \\ 0 & \varepsilon^{-1} \end{pmatrix}$$

erzeugt. Ein Erzeugendensystem für T kann dann in folgender Weise berechnet werden. Da T in M den Index 10 hat, gibt es eine Zerlegung

$$M = \sum_{i=1}^{10} T G_i \quad (G_1 = E).$$

Liegt eine Substitution $G \in M$ in der Restklasse $T G_i$, so sei $G = G_i$. Die 30 Substitutionen

$$U_{ik} = G_i S_k \overline{G_i S_k}^{-1} \quad (i = 1, \dots, 10; k = 1, 2, 3)$$

erzeugen dann die Thetagruppe T . Wählen wir für G_i ($i = 1, \dots, 10$) der Reihe nach

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \varepsilon \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1+\varepsilon \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \varepsilon & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1+\varepsilon & 1 \end{pmatrix} \\ \begin{pmatrix} 1+\varepsilon & \varepsilon \\ -\varepsilon & 1-\varepsilon \end{pmatrix} \begin{pmatrix} 1 & 1 \\ \varepsilon-1 & \varepsilon \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -\varepsilon & 1-\varepsilon \end{pmatrix},$$

so bestätigt man leicht, daß zwischen den U_{ik} die nachfolgenden 27 Relationen erfüllt sind:

$$\begin{aligned} U_{11} &= U_{31} = E, \\ U_{41} &= U_{21}, \\ U_{52} &= U_{51} = U_{12} U_{21}^{-1}, \\ U_{61} &= U_{12} U_{51}^{-1} U_{13}^{-1} U_{51} U_{12}^{-1} U_{13}, \\ U_{71} &= U_{51} U_{12}^{-1}, \\ U_{81} &= U_{21}^{-1} U_{13}^{-2} U_{71} U_{61}^{-1} U_{12}, \\ U_{22} &= U_{21} U_{12}, \\ U_{32} &= U_{12} U_{61}^{-1}, \\ U_{102} &= U_{101} = U_{12} U_{32}^{-1} U_{13}^3 U_{71} U_{61}^2 U_{12}^{-1} U_{71}^{-1} U_{61} U_{12}, \end{aligned}$$

⁷⁾ F. Götzky, Über eine zahlentheoretische Anwendung von Modulfunktionen zweier Veränderlicher. Math. Annalen 100 (1928), S. 411–437.

$$\begin{aligned}
U_{91} &= U_{101}^{-1} U_{12}^2 U_{61}^{-1}, \\
U_{42} &= U_{22} U_{12}^{-1} U_{32}, \\
U_{62} &= U_{61} U_{12}, \\
U_{72} &= U_{12}^2 U_{42}^{-1}, \\
U_{82} &= U_{21}^{-1} U_{42}^{-1} U_{71}^2 U_{61}^{-1}, \\
U_{92} &= U_{12} U_{101} U_{12}, \\
U_{23} &= U_{13} U_{51}^{-1} U_{62}, \\
U_{33} &= U_{13} U_{21}^{-1}, \\
U_{43} &= U_{53} = U_{13}, \\
U_{63} &= U_{13} U_{61}, \\
U_{73} &= U_{71} U_{13}, \\
U_{83} &= U_{12}^2 U_{13}^3 U_{71}^{-2} U_{61}^{-3} U_{12}^{-1} U_{71}^{-2} U_{61} U_{12} U_{61}^{-1}, \\
U_{93} &= U_{12}^{-1}, \\
U_{103} &= U_{61}^{-1} U_{12}^{-1}.
\end{aligned}$$

Demnach wird \mathbf{T} von den drei Substitutionen U_{21} , U_{12} , U_{13} erzeugt. In gleicher Reihenfolge sind das die Substitutionen

$$(12) \quad U_1 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad U_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad U_3 = \begin{pmatrix} \varepsilon & 0 \\ 0 & \varepsilon^{-1} \end{pmatrix}.$$

Wir beginnen mit der Untersuchung von \mathbf{M} und denken uns ein beliebiges Multiplikatorsystem v von \mathbf{M} zur Dimension $-r$ gegeben. Es soll zunächst

$$(13) \quad v\left(\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}\right) = 1 \quad \text{für } \alpha < 0$$

gezeigt werden. Dazu bestimme man (vgl.¹⁾) den Modul t der Translationen in der affinen Gruppe von \mathbf{M} , ferner den Modul m als Gesamtheit der Lösungen μ von

$$S \frac{\mu t}{\sqrt{5}} \equiv 0[1]$$

und schließlich, wenn

$$v\left(\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}\right) = e^{2\pi i \varrho(\alpha)} \quad \text{für } \alpha < t,$$

α aus

$$S \frac{\alpha}{\sqrt{5}} \equiv \varrho(\alpha)[1] \quad \text{für alle } \alpha < t.$$

Da ε^2 unter den Multiplikatoren der affinen Substitutionen aus \mathbf{M} vorkommt, so gilt

$$\alpha(1 - \varepsilon^2) = -\varepsilon \alpha < m.$$

Offenbar ist nun $t = [1, \varepsilon]$, $m = [1, \varepsilon]$, woraus folgt, daß $\alpha < m$ und $\varrho(\alpha) \equiv 0[1]$, womit (13) bewiesen ist. Aus den Relationen

$$\begin{pmatrix} 0 & \varepsilon \\ 1-\varepsilon & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & \varepsilon \\ 1-\varepsilon & 1 \end{pmatrix}^3 = -E, \quad \begin{pmatrix} 0 & \varepsilon \\ 1-\varepsilon & 1 \end{pmatrix} = \begin{pmatrix} 0 & \varepsilon \\ 1-\varepsilon & 0 \end{pmatrix} \begin{pmatrix} 1 & -\varepsilon \\ 0 & 1 \end{pmatrix}$$

wird auf Grund von (2) und (3)

$$v \left(\begin{pmatrix} 0 & \varepsilon \\ 1-\varepsilon & 0 \end{pmatrix} \right) = e^{2\pi i \left(\frac{r}{2} + \frac{a}{2} \right)}, \quad v \left(\begin{pmatrix} 0 & \varepsilon \\ 1-\varepsilon & 1 \end{pmatrix} \right) = e^{2\pi i \left(\frac{r}{3} + \frac{b}{3} \right)}$$

$$v \left(\begin{pmatrix} 0 & \varepsilon \\ 1-\varepsilon & 1 \end{pmatrix} \right) = v \left(\begin{pmatrix} 0 & \varepsilon \\ 1-\varepsilon & 0 \end{pmatrix} \right) v \left(\begin{pmatrix} 1 & -\varepsilon \\ 0 & 1 \end{pmatrix} \right)$$

erschlossen, wobei $a = 0$ oder 1 und b aus der Reihe $0, 1, 2$. Zusammen mit (13) folgt nun

$$r \equiv 3a + 2b \pmod{6},$$

d. h. r ist ganzzahlig.

Die analoge Betrachtung ist für T anzustellen. Die Maximalzahl inäquivalenter parabolischer Spitzen von T beträgt 2; wir wählen als Repräsentanten $\infty, -\varepsilon$. Durch

$$A = \begin{pmatrix} \varepsilon & \varepsilon \\ 1 & \varepsilon \end{pmatrix}$$

wird die zweite Spitze nach ∞ transformiert. Mit t bzw. t_A bezeichnen wir die Moduln der Translationen in den affinen Gruppen von T bzw. $AT A^{-1}$ und bestimmen m bzw. m_A als Gesamtheit der Lösungen μ von

$$S \frac{\mu t}{2\sqrt{5}} \equiv 0 [1] \quad \text{bzw.} \quad S \frac{\mu t_A}{2\sqrt{5}} \equiv 0 [1].$$

Setzt man noch für ein vorgegebenes Multiplikatorsystem v von T zur Dimension $-\tau$

$$v \left(\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \right) = e^{2\pi i \varrho(\alpha)} \quad \text{für } \alpha \in t$$

und für das mit A^{-1} transformierte Multiplikatorsystem $v A^{-1}$ analog

$$v A^{-1} \left(\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \right) = e^{2\pi i \varrho_A(\alpha)} \quad \text{für } \alpha \in t_A,$$

so sind κ bzw. κ_A wieder durch

$$S \frac{\kappa \alpha}{2\sqrt{5}} \equiv \varrho(\alpha) [1] \quad \text{für alle } \alpha \in t$$

bzw.

$$S \frac{\kappa_A \alpha}{2\sqrt{5}} \equiv \varrho_A(\alpha) [1] \quad \text{für alle } \alpha \in t_A$$

zu bestimmen. ε^2 kommt unter den Multiplikatoren der affinen Substitutionen von T vor, dagegen erst ε^6 unter den Multiplikatoren der affinen Substitutionen von $AT A^{-1}$. Folglich gilt

$$\kappa(1 - \varepsilon^2) = -\varepsilon \kappa \in m,$$

$$\kappa_A(1 - \varepsilon^6) = -4\varepsilon^3 \kappa_A \in m_A.$$

Man findet nun leicht:

$t = [2, 2\varepsilon]$, $m = [1, \varepsilon]$, $t_A = [2, \varepsilon - 1]$, $m_A = [2, \varepsilon]$
und damit

$$x \subset m, \quad 4x_A \subset m_A,$$

insbesondere also

$$(14) \quad \varrho(x) \equiv 0[1] \text{ für } x \subset t, \quad 4\varrho_A(x) \equiv 0[1] \text{ für } x \subset t_A.$$

Aus (2), angewendet auf

$$\begin{pmatrix} 2 & \varepsilon \\ 1-\varepsilon & 0 \end{pmatrix} = \begin{pmatrix} 1 & -2\varepsilon \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & \varepsilon \\ 1-\varepsilon & 0 \end{pmatrix} \\ \begin{pmatrix} 2 & \varepsilon \\ 1-\varepsilon & 0 \end{pmatrix} = A^{-1} \begin{pmatrix} 1 & \varepsilon-1 \\ 0 & 1 \end{pmatrix} A,$$

folgt

$$v \left(\begin{pmatrix} 2 & \varepsilon \\ 1-\varepsilon & 0 \end{pmatrix} \right) = v \left(\begin{pmatrix} 1 & -2\varepsilon \\ 0 & 1 \end{pmatrix} \right) v \left(\begin{pmatrix} 0 & \varepsilon \\ 1-\varepsilon & 0 \end{pmatrix} \right).$$

Wie eine kleine Rechnung zeigt, ist

$$v \left(\begin{pmatrix} 2 & \varepsilon \\ 1-\varepsilon & 0 \end{pmatrix} \right) = v A^{-1} \left(\begin{pmatrix} 1 & \varepsilon-1 \\ 0 & 1 \end{pmatrix} \right)$$

und wie oben

$$v \left(\begin{pmatrix} 0 & \varepsilon \\ 1-\varepsilon & 0 \end{pmatrix} \right) = e^{2\pi i \left(\frac{r}{2} + \frac{a}{2} \right)} \quad (\alpha = 0 \text{ oder } 1).$$

Damit ergibt sich

$$\varrho_A(\varepsilon - 1) \equiv \frac{r}{2} + \frac{a}{2} + \varrho(-2\varepsilon)[1],$$

wegen (14) also

$$2r \equiv 0[1],$$

d. h. r ist entweder ganz- oder halbzahlig.

Wir kommen nun zur Aufstellung der Multiplikatorsysteme und betrachten zunächst wieder M . Jedes Multiplikatorsystem zu M und zu ganzzahliger Dimension liefert offenbar eine Darstellung der Faktorgruppe von M nach der Kommutatorgruppe K . Es kommt also darauf an, diese zu bestimmen. Aus den Relationen

$$(15) \quad S_2 S_1^{-1} S_2 (S_2 S_1^{-1} S_3 S_2 S_1^{-1})^2 = S_3 S_2 S_1^{-1} S_3, \\ S_2^2 = (S_2 S_1^{-1})^3 = -E, (S_3 S_2 S_1^{-1})^5 = E, S_3 S_2 S_3 = S_2,$$

die man leicht verifiziert, ergibt sich, wenn man mit den Erzeugenden S_i abelsch rechnet, daß alle drei Erzeugende in K enthalten sind. Damit ist

$$(16) \quad M = K$$

gezeigt und bewiesen, daß es nur das eine Multiplikatorsystem $v(S) = 1$ für $S \subset M$ gibt.

Ähnliche Überlegungen sind für T anzustellen. v sei ein Multiplikatorsystem zu T und zu ganzzahliger Dimension. Da notwendig $v(-E) = 1$, so reicht es, in T die Gruppe K_0 zu bestimmen, die aus den Kommutatoren von T und $-E$ erzeugt wird. Offenbar liegen die Substitutionen

$$\begin{aligned} -E &= U_2^2, & U_2 U_3^{-1} U_2^{-1} U_3 &= U_3^2, \\ U_3 U_1 U_3^{-1} U_1^{-1} &= V, & V U_3 V^{-1} U_3^{-1} &= U_1 \end{aligned}$$

in K_0 . Unter den Substitutionen

$$E, U_2, U_3, U_2 U_3$$

gibt es also ein vollständiges Restsystem mod K_0 . Bezeichnen wir mit $M(2)$ die Hauptkongruenzuntergruppe von M zur Stufe 2, so hat man in

$$N = \sum_{r=0}^2 M(2) U_3^r$$

einen Normalteiler von T vom Index 2, und es gilt die Zerlegung

$$T = N + N U_2.$$

Durch

$$(17) \quad v_1(S) = 1, \quad v_1(S U_2) = -1 \quad \text{für } S \in N$$

ist dann ein Multiplikatorsystem zu ganzzahliger Dimension definiert. Ein zweites davon unabhängiges hat man in $v_{0\infty}^2$; dabei bedeutet $v_{0\infty}$ das ϑ -Multiplikatorsystem zu T^3). Für dieses ist

$$(18) \quad v_{0\infty}(U_1) = v_{0\infty}(U_2) = 1, \quad v_{0\infty}(U_3) = i.$$

Damit hat man die vier Multiplikatorsysteme

$$(19) \quad v_{0\infty}^{2a} \cdot v_1^b \quad (a = 0, 1; b = 0, 1)$$

zu ganzzahliger Dimension gewonnen. Andere kann es nicht geben, da ja K_0 in T den Index 4 hat. Sei v ein Multiplikatorsystem zu T und zu halbzahliger Dimension, dann ist $v v_{0\infty}^{-1}$ ein solches zu ganzzahliger Dimension. Es gibt also auch genau vier Multiplikatorsysteme zu T und zu halbzahliger Dimension, nämlich

$$(20) \quad v_{0\infty}^{2a+1} \cdot v_1^b \quad (a = 0, 1; b = 0, 1).$$

Die Werte von $v_{0\infty}$ sind in ³⁾ explizit berechnet. Wir fassen das Ergebnis zusammen in

Satz 1. 1. Für M gibt es nur zu ganzzahligen Dimensionen ein Multiplikatorsystem, und zwar genau eins, nämlich das triviale.

2. Für T gibt es nur zu ganz- und halbzahligen Dimensionen ein Multiplikatorsystem, und zwar in jedem Fall genau vier. Sie sind durch (19) und (20) gegeben.

§ 2.

Modulformen zu M .

Eine im Bereich (1) reguläre Funktion $\varphi_r(\tau) = \varphi_r(\tau, \tau')$ mit der Transformationseigenschaft

$$(21) \quad \varphi_r(S\tau) = N(\gamma\tau + \delta)^r \varphi_r(\tau) \quad \text{für } S \in M,$$

wobei $S = (\gamma, \delta)$ die zweite Zeile von S , heißt eine Modulform zu M und zur (ganzahligen) Dimension $-r$, wenn $\varphi_r(\tau)$ auch im Unendlichen regulär ist, d. h. wenn eine Entwicklung der Art

$$(22) \quad \varphi_r(\tau) = \sum_{\substack{\nu \in \mathbb{Z} \\ \nu \geq 0}} b_r(\nu) e^{2\pi i S \frac{\nu}{5}}$$

gilt. Werden die unabhängigen Veränderlichen τ und τ' durch die Gleichung

$$\tau_1 = \varepsilon\tau = \varepsilon'\tau'$$

gekoppelt, so geht $\varphi_r(\tau)$ in eine Modulform

$$(23) \quad f_{2r}(\tau_1) = \sum_{m=0}^{\infty} a_{2r}(m) e^{2\pi i \tau_1 m}$$

zur rationalen Modulgruppe M_0 und zur Dimension $-2r$ mit den Koeffizienten

$$(24) \quad a_{2r}(m) = \sum_{\substack{\nu = m + n\varepsilon \\ \nu \geq 0}} b_r(\nu)$$

über. Die ersten Koeffizienten lauten demnach

$$(25) \quad \begin{aligned} a_{2r}(0) &= b_r(0), \\ a_{2r}(1) &= b_r(1) + b_r(1 + \varepsilon), \\ a_{2r}(2) &= b_r(2 - \varepsilon) + b_r(2) + b_r(2 + \varepsilon) + b_r(2 + 2\varepsilon) + b_r(2 + 3\varepsilon). \end{aligned}$$

Wendet man (21) für $S = S_3$ an, so erhält man

$$\varphi_r(\varepsilon^2\tau) = (-1)^r \varphi_r(\tau),$$

und nach (22)

$$(26) \quad b_r(\varepsilon^2\nu) = (-1)^r b_r(\nu).$$

Je nachdem r gerade oder ungerade, so gilt also

$$(27) \quad \begin{aligned} a_{2r}(0) &= b_r(0), \\ a_{2r}(1) &= 2b_r(1), \\ a_{2r}(2) &= 2(b_r(1) + b_r(2)) + b_r(2 + \varepsilon) \end{aligned} \quad (r \equiv 0 \pmod{2})$$

oder

$$(28) \quad \begin{aligned} a_{2r}(0) &= b_r(0) = 0, \\ a_{2r}(1) &= 0, \\ a_{2r}(2) &= b_r(2 + \varepsilon). \end{aligned} \quad (r \equiv 1 \pmod{2})$$

Durch Quadrieren von $\varphi_r(\tau)$ gelangen wir schließlich zu einer Form

$$(29) \quad \varphi_{2r}(\tau) = \sum_{\substack{r_1 + r_2 = r \\ r_i \geq 0}} b_{2r}(\nu) e^{\frac{2\pi i S \nu \tau}{\sqrt{5}}}$$

von der Dimension $-2r$ mit den Koeffizienten

$$(30) \quad b_{2r}(\nu) = \sum_{\substack{r_1 + r_2 = r \\ r_i \geq 0}} b_r(\nu_1) b_r(\nu_2).$$

Speziell von

$$(31) \quad \begin{aligned} b_{2r}(0) &= (b_r(0))^2, \\ b_{2r}(2) &= (b_r(1))^2 + 2 b_r(0) b_r(2), \\ b_{2r}(3) &= 2 (b_r(0) b_r(3) + b_r(1) b_r(2) + (b_r(1))^2) \end{aligned}$$

soll später Gebrauch gemacht werden.

Wir wollen jetzt voraussetzen, daß in der Entwicklung (22) alle Koeffizienten zu Exponenten mit einer Norm unterhalb einer positiven Schranke T verschwinden:

$$(32) \quad b_r(\nu) = 0 \text{ sobald } N \nu \leq T \quad (T > 0).$$

Die Schranke T soll möglichst klein so bestimmt werden, daß aus (32) das identische Verschwinden von $\varphi_r(\tau)$ erschlossen werden kann. Für die Durchführung dieses Schlusses ist es entscheidend, daß es für \mathbf{M} einen Fundamentalbereich \mathfrak{F} im Bereich (1) mit folgenden Eigenschaften gibt⁷⁾. Wir setzen

$$(33) \quad \tau = s + it, \quad \tau' = s' - it'.$$

Der Bereich (1) wird dann beschrieben durch die Ungleichungen

$$(34) \quad t > 0, \quad t' > 0.$$

Für jeden Punkt des Fundamentalbereiches \mathfrak{F} gilt

$$(35) \quad t \geq 0,45, \quad t' \geq 0,45,$$

und nähert man sich im Fundamentalbereich \mathfrak{F} an den Rand des Bereiches (34), so gehen t und t' gleichzeitig gegen ∞ . Auf Grund der Voraussetzung (32)

liegt das Maximum M von $(t t')^{\frac{r}{2}} |\varphi_r(\tau)|$ für $\{\tau, \tau'\} \subset \mathfrak{F}$ in einem endlichen Punkt

$$\tau_0 = s_0 + it_0, \quad \tau'_0 = s_0 - it_0$$

von \mathfrak{F} . Da nun $(t t')^{\frac{r}{2}} |\varphi_r(\tau)|$ gegenüber den Substitutionen aus \mathbf{M} invariant ist, so gilt überhaupt

$$(36) \quad (t t')^{\frac{r}{2}} |\varphi_r(\tau)| \leq M.$$

In der Integraldarstellung

$$(37) \quad b_r(\nu) = \frac{1}{\Delta} \iint_{\mathfrak{P}} \varphi_r(\tau) e^{-2\pi i S \frac{\nu \tau}{\sqrt{5}}} ds ds', \quad \Delta = \iint_{\mathfrak{P}} ds ds'$$

(\mathfrak{P} = Grundmasche des Gitters der Translationen aus \mathbf{M}) setze man

$$t = (1 - \theta) t_0, \quad t' = (1 - \theta) t'_0, \quad 0 < \theta < 1$$

und erhält damit unter Berücksichtigung von (36) aus (37) die Abschätzung

$$|b_r(\nu)| \leq M (1 - \theta)^{-r} (t_0 t'_0)^{-\frac{r}{2}} e^{-\frac{r}{2} \frac{2\pi(1-\theta)}{\sqrt{5}} \frac{\nu t_0 + \nu' t'_0}{\sqrt{5}}}$$

$\varphi(\tau_0)$ wird jetzt erneut nach (22) und (32) abgeschätzt mit dem Resultat

$$M = (t_0 t'_0)^{\frac{r}{2}} |\varphi_r(\tau_0)| \leq M (1 - \theta)^{-r} \sum_{\substack{N\nu > T \\ \nu \geq 0}} e^{-2\pi \theta \frac{\nu t_0 + \nu' t'_0}{\sqrt{5}}}.$$

Um daraus zu schließen, daß M verschwindet, genügt es wegen $t_0, t'_0 \geq 0,45$, die Bedingung

$$(38) \quad F(T) = \sum_{\substack{N\nu > T \\ \nu \geq 0}} e^{-\frac{9\pi\theta}{10\sqrt{5}} S\nu} < (1 - \theta)^r$$

durch geeignete Wahl von T zu erreichen. Setzen wir

$$(39) \quad h = e^{-\frac{9\pi\theta}{10\sqrt{5}}}$$

und bezeichnen wir mit $B_T(k)$ die Anzahl der Zahlen $\nu < 0$, für welche

$$\nu > 0, \quad N\nu > T, \quad S\nu = k,$$

so ist

$$(40) \quad F(T) = \sum_{k=1}^{\infty} B_T(k) h^k.$$

Macht man einen Ansatz $\nu = m + n\varepsilon$, dann ist $k = S\nu = 2m + n$, $4N\nu = k^2 - 5n^2$ und daher $B_T(k)$ gleich der Anzahl der ganz rationalen Zahlen n , welche den Bedingungen

$$k^2 - 4T > 5n^2, \quad k \equiv n(2)$$

genügen, woraus insbesondere

$$(41) \quad B_T(k) < \frac{k}{\sqrt{5}} + 1$$

folgt. Benutzt man diese Abschätzung für alle $k > 12$, so erhalten wir nach (40)

$$F(T) < \sum_{k=1}^{12} B_T(k) h^k + \frac{1}{\sqrt{5}} \left[\frac{13h^{13}}{1-h} + \frac{h^{14}}{(1-h)^2} \right] + \frac{h^{15}}{1-h}.$$

Wir wählen jetzt $h = \frac{1}{2}$, bestimmen durch direkte Abzählung die folgenden Werte:

k	1	2	3	4	5	6	7	8	9	10	11	12
$B_5(k)$	0	0	0	0	0	1	2	3	4	3	4	5
$B_9(k)$	0	0	0	0	0	0	2	3	2	3	4	5

und stellen alsdann fest, daß

$$F(5) < 0,06, \quad F(9) < 0,040.$$

θ ist nach Wahl von h durch (39) bestimmt:

$$\theta = \frac{10 \cdot \sqrt{5} \cdot \log 2}{9 \cdot \pi}.$$

Für diesen Wert ist

$$0,09 < (1 - \theta)^3, \quad 0,041 < (1 - \theta)^4,$$

so daß also

$$F(5) < (1 - \theta)^3, \quad F(9) < (1 - \theta)^4.$$

Für die Dimension -4 besagt diese Rechnung, daß $\varphi_4(\tau)$ identisch verschwindet, sobald $b_4(v) = 0$ für $v \geq 0$, $Nv \leq 9$. Nach (26) genügt es, wenn v die Werte $0, 1, 2, 2 + \varepsilon, 3$ durchläuft. Wenn $b_4(0) = 0$ so ist in (27) $a_8(n) = 0$ für alle n , da es keine Spitzenform zur rationalen Modulgruppe M_0 von der Dimension -8 gibt, also ist auch $b_4(1) = 0$. Aus $b_4(2) = 0$ folgt dann nach (27) auch $b_4(2 + \varepsilon) = 0$. Man braucht also nur das Verschwinden von drei Koeffizienten zu fordern.

Im Falle der Dimension -3 muß $b_3(v) = 0$ für $v = 0, 1, 2, 2 + \varepsilon$ gefordert werden. Analog wie oben schließt man, daß in (28) $a_6(n) = 0$ für alle n gilt. $b_3(0) = b_3(2 + \varepsilon) = 0$ ist also automatisch erfüllt, und es bleibt nur noch, $b_3(1) = b_3(2) = 0$ zu fordern. $\varphi_2(\tau)$ verschwindet bereits dann, wenn $b_2(0) = 0$. Nach (27) ist nämlich wie oben zu schließen, daß $b_2(1) = 0$. Dann verschwinden nach (31) auch die Koeffizienten $b_4(v)$ für $v = 0, 2, 3$, und das reicht ja aus, um $(\varphi_2(\tau))^2 \equiv 0$ zu beweisen. Eine Form $\varphi_1(\tau)$ gibt es nicht; denn nach (28) ist $b_1(0) = 0$, nach (31) somit $b_2(0) = 0$ und daher, wie soeben bewiesen wurde, $(\varphi_1(\tau))^2 \equiv 0$. Wir formulieren dieses Ergebnis in

Satz 2. Eine Modulform

$$\varphi_r(\tau) = \sum_{\substack{r \subseteq s \\ r \geq 0}} b_r(v) e^{\frac{2\pi i s}{\sqrt{5}}}$$

zu M und zur Dimension $-r$ verschwindet identisch, wenn

1. $b_4(0) = b_4(2) = b_4(3) = 0$ für $r = 4$,
2. $b_3(1) = b_3(2) = 0$ für $r = 3$,
3. $b_2(0) = 0$ für $r = 2$,
4. stets für $r = 1$.

Beispiele von Modulformen zu \mathbf{M} sind gegeben in den Eisensteinreihen

$$(42) \quad G_r(\tau) = 1 + \sum_{\substack{r \leq \frac{r}{2} \\ r \geq 0}} C_r(\nu) e^{2\pi i S \frac{\tau}{\sqrt{5}}}$$

zu gerader Dimension $-r^0$ mit den Koeffizienten

$$(43) \quad C_r(\nu) = \frac{(2\pi)^{2r} \cdot \sqrt{5}}{(r(\tau))^2 \cdot 5^r \cdot \zeta_Z(\tau)} \cdot \sum_{(u)/r} |N\mu|^{r-1},$$

wobei ζ_Z die Zetafunktion zum Zahlkörper $Z = R(\sqrt{5})$. Der Zahlfaktor vor dem Summenzeichen auf der rechten Seite der Gleichung (43) hat für $r = 2$ bzw. 4 die Werte 120 bzw. 240. Speziell findet man

$$(44) \quad \begin{aligned} C_2(1) &= 120, & C_2(2) &= 600, & C_2(3) &= 1200, \\ C_4(2) &= 15600, & C_4(3) &= 175200. \end{aligned}$$

Berechnet man nach (31) die Koeffizienten von $(G_2(\tau))^2$ zu den Exponenten 2 und 3, so ergibt sich auf Grund von Satz 2 die Identität $(G_2(\tau))^2 = G_4(\tau)$.

Wir betrachten jetzt die eingangs genannten Thetareihen $\vartheta(\tau, \Omega_m)$ zu Klassen Ω_m positiv definiter ganzzahliger gerader quadratischer Formen $Q(x_1, \dots, x_m)$ von m Veränderlichen und mit Determinante ε^{2r} . Für diese Reihen gelten die drei Transformationsgleichungen

$$\vartheta(S_1\tau, \Omega_m) = \vartheta(S_3\tau, \Omega_m) = \vartheta(\tau, \Omega_m),$$

$$\vartheta(S_2\tau, \Omega_m) = N(-\tau)^{\frac{m}{2}} \vartheta(\tau, \Omega_m).^8)$$

$\vartheta(\tau, \Omega_m)$ ist also eine Modulform zu \mathbf{M} von der Dimension $-\frac{m}{2}$. Da es zu \mathbf{M} nur das triviale Multiplikatorsystem gibt, so ist notwendig

$$\vartheta(S_3\tau, \Omega_m) = N\varepsilon^{-\frac{m}{2}} \vartheta(\tau, \Omega_m) = \vartheta(\tau, \Omega_m),$$

also $m \equiv 0(4)$. Zu $m = 4$ gibt es die Form

$$(45) \quad \begin{pmatrix} 2 & -1 & 0 & 1-\varepsilon \\ -1 & 2 & -1 & \varepsilon-1 \\ 0 & -1 & 2 & -\varepsilon \\ 1-\varepsilon & \varepsilon-1 & -\varepsilon & 2 \end{pmatrix}.$$

Ω_4 sei die Klasse, welche diese Form enthält. Durch Bildung direkter Summen erhält man dann quadratische Formen der genannten Art zu jeder durch 4 teilbaren Veränderlichenzahl. Nach Satz 2 ist $\vartheta(\tau, \Omega_4) = G_2(\tau)$. Unter Ω_8 soll die Formenklasse verstanden werden, welche die rationale Klasse Ω_8^*

⁸⁾ H. D. Kloosterman, Thetareihen in total-reellen algebraischen Zahlkörpern. Math. Annalen 103 (1930), S. 279–299.

der geraden Formen mit ganz rationalen Koeffizienten enthält. Eine Form der rationalen Klasse Ω_8^* gestattet folgende Darstellung⁶⁾: Man betrachte das Gitter Ξ_8 aller Vektoren $x = (\xi_1, \dots, \xi_8)$ mit rationalzahligen Komponenten, für welche

$$\xi_i \equiv 0 \left(\frac{1}{2}\right), \quad \xi_i \equiv \xi_k \pmod{1}, \quad \sum_{i=1}^8 \xi_i \equiv 0 \pmod{2},$$

und bezeichne mit u_1, \dots, u_8 eine Gitterbasis. Dann ist

$$\left(\sum_{i=1}^8 x_i u_i\right)^2 = \sum_{i,k=1}^8 (u_i u_k) x_i x_k$$

eine ganzzahlige gerade Form mit Determinante 1. Sie ist daher in Ω_8^* enthalten. Mit e_1, \dots, e_8 bezeichnen wir die Einheitsvektoren des 8-dimensionalen euklidischen Raumes. Wie man sich leicht überlegt, sind dann die Vektoren $a \in \Xi_8$ mit $a^2 = 2$ bzw. 4 von der Gestalt

$$(46) \quad \pm e_i \pm e_k \quad (i \neq k), \quad \frac{1}{2} \sum_{i=1}^8 \varepsilon_i e_i \quad (\varepsilon_i = \pm 1, \prod_{i=1}^8 \varepsilon_i = 1),$$

bzw.

$$(47) \quad \pm 2e_i, \quad \pm e_i \pm e_k \pm e_l \pm e_m \quad (i < k < l < m) \\ \frac{1}{2} \left(\sum_{i=1}^8 \varepsilon_i e_i \right) + e_k e_l \quad (\varepsilon_i = \pm 1, \prod_{i=1}^8 \varepsilon_i = -1).$$

Es soll jetzt $\vartheta(\tau, \Omega_8) = G_4(\tau)$ bewiesen werden. Bedeutet allgemein $a(2\nu, \Omega_m)$ die Anzahl der ganzzahligen Darstellungen der geraden Zahl 2ν durch eine Form der Klasse Ω_m , so daß also

$$(48) \quad \vartheta(\tau, \Omega_m) = 1 + \sum_{\substack{\nu \in \mathbb{Z} \\ \nu > 0}} a(2\nu, \Omega_m) e^{2\pi i \nu \frac{\tau}{\sqrt{6}}},$$

so brauchen wir nach (44) und Satz 2 nur

$$(49) \quad a(4, \Omega_8) = 15600, \quad a(6, \Omega_8) = 175200$$

zu bestätigen. Diese beiden Darstellungsanzahlen werden jetzt durch direkte Abzählung in folgender Weise bestimmt. $a_0(2n, \Omega_8^*)$ sei die Anzahl der ganzzahligen Darstellungen der geraden Zahl $2n$ durch eine Form der Klasse Ω_8^* . Die Thetareihe

$$\vartheta_0(\tau_1, \Omega_8^*) = 1 + \sum_{n=1}^{\infty} a_0(2n, \Omega_8^*) e^{2\pi i \tau_1 n}$$

ist dann bekanntlich mit der normierten Eisensteinreihe zur rationalen Modulgruppe M_0 und zur Dimension -4 identisch. Aus dieser Identität werden die Anzahlen

$$(50) \quad a_0(4, \Omega_8^*) = 2160, \quad a_0(6, \Omega_8^*) = 6720$$

abgeleitet⁶⁾. Wir betrachten nun die Vektoren $x = \sum_{i=1}^8 x_i u_i$ mit $x_i \in \mathfrak{o}$. Da $1, \varepsilon$ eine Basis von \mathfrak{o} , so erhält man offenbar jeden solchen Vektor, wenn in $x = a + b\varepsilon$ die Vektoren a, b unabhängig alle Gittervektoren aus Ξ_8 durchlaufen. Es ist dann

$$x^2 = a^2 + b^2 + (2ab + b^2)\varepsilon.$$

Für $x^2 = 4$ bzw. 6 muß daher

$$a^2 + b^2 = 4, \quad ab = -\frac{1}{2}b^2$$

bzw.

$$a^2 + b^2 = 6, \quad ab = -\frac{1}{2}b^2$$

zutreffen. Die Abzählung der Lösungspaare a, b wird so vorgenommen, daß man alle Paare a, b mit gleichem Wert a^2 zu einem Komplex zusammenfaßt. Diesem Prozeß entspricht die Zerlegung

$$(51) \quad \begin{aligned} a(4, \Omega_8) &= a_0(4, \Omega_8^*) + a^*, \\ a(6, \Omega_8) &= a_0(6, \Omega_8^*) + a_1^* + a_2^*, \end{aligned}$$

wobei a^* gleich der Anzahl der Paare $a, b \in \Xi_8$ mit

$$a^2 = b^2 = 2, \quad ab = -1,$$

a_1^* gleich der Anzahl der Paare $a, b \in \Xi_8$ mit

$$a^2 = 4, \quad b^2 = 2, \quad ab = -1$$

und a_2^* gleich der Anzahl der Paare $a, b \in \Xi_8$ mit

$$a^2 = 4, \quad b^2 = 2, \quad ab = -2.$$

Beachtet man, daß die vorkommenden Vektoren a, b sämtlich durch (46) und (47) gegeben sind, und diskutiert alle möglichen Kombinationen, so erhält man

$$a^* = 13440, \quad a_1^* = 138240, \quad a_2^* = 30240,$$

woraus sich mit (51) und (50) wirklich (49) ergibt. Wir erhalten folgendes Resultat:

Satz 3. *Es bestehen die Identitäten*

$$\vartheta(\tau, \Omega_4) = G_2(\tau), \quad \vartheta(\tau, \Omega_8) = G_4(\tau), \quad (G_2(\tau))^2 = G_4(\tau),$$

d. h. es ist

$$\begin{aligned} a(2\nu, \Omega_4) &= 120 \cdot \sum_{(u)\tau} |N\mu|, \\ a(2\nu, \Omega_4 + \Omega_4) &= a(2\nu, \Omega_8) = 240 \cdot \sum_{(u)\tau} |N\mu|^2. \end{aligned}$$

§ 3.

Bestimmung von Klassen quadratischer Formen.

Das Ziel der folgenden Betrachtungen ist der Nachweis, daß es außer den Klassen Ω_4 , $\Omega_4 + \Omega_4$, Ω_8 keine anderen Klassen gerader Formen mit Determinante ε^{2r} in 4 oder 8 Veränderlichen mehr gibt. Zu diesem Zweck sollen die Gitter, welche den Klassen solcher Formen zugeordnet werden, einer näheren Betrachtung unterzogen werden. Wir betrachten eine beliebige ganzzahlige gerade Form $Q(x_1 \dots x_m)$ über $R(\sqrt{5})$ mit Determinante ε^{2r} und denken uns Vektoren u_1, \dots, u_m derart bestimmt, daß

$$(52) \quad \mathbf{x}^2 = Q(x_1, \dots, x_m) \text{ mit } \mathbf{x} = \sum_{i=1}^m x_i u_i$$

gilt. In dem Gitter U_m aller Vektoren $\mathbf{x} = \sum_{i=1}^m x_i u_i$ mit Koeffizienten $x_i \in \mathfrak{o}$ ist die Teilmenge V_m der Vektoren \mathbf{x} mit $\mathbf{x}^2 = 2$ von besonderem Interesse. Für zwei beliebige Vektoren $\mathbf{x}_1, \mathbf{x}_2 \in U_m$ ist stets $\mathbf{x}_1 \mathbf{x}_2 \in \mathfrak{o}$; denn es ist

$$(\mathbf{x}_1 + \mathbf{x}_2)^2 - \mathbf{x}_1^2 - \mathbf{x}_2^2 = 2 \mathbf{x}_1 \mathbf{x}_2 \in 2 \mathfrak{o}.$$

Für Zahlen α aus $R(\sqrt{5})$ sei durch $\alpha \rightarrow \alpha'$ der Automorphismus des Körpers bezeichnet. Man bestimme nun Vektoren u'_1, \dots, u'_m derart, daß

$$\left(\sum_{i=1}^m x_i u'_i \right)^2 = Q'(x_1, \dots, x_m)$$

mit der zu $Q(x_1, \dots, x_m)$ konjugierten Form übereinstimmt; setzt man noch $\mathbf{x}' = \sum_{i=1}^m x'_i u'_i$ für $\mathbf{x} \in U_m$, so ist $\mathbf{x}'_1 \mathbf{x}'_2$ für $\mathbf{x}_1, \mathbf{x}_2 \in U_m$ in gewöhnlichem Sinne zu $\mathbf{x}_1 \mathbf{x}_2$ konjugiert, d. h. es ist

$$(\mathbf{x}_1 \mathbf{x}_2)' = \mathbf{x}'_1 \mathbf{x}'_2.$$

Angewendet auf Vektoren $a_1, a_2 \in V_m$ folgt daraus, weil

$$a_i^2 = a'_i{}^2 = 2 \quad (i = 1, 2) \quad \text{und} \quad |a_1 a_2| \leq 2, \quad |a'_1 a'_2| \leq 2,$$

daß die ganze Zahl $a_1 a_2$ sowie ihre Konjugierte dem Betrage nach höchstens gleich 2 sind. Daher kommen nur folgende Skalarprodukte in Betracht:

$$a_1 a_2 = 2, \varepsilon, 1, \varepsilon^{-1}, 0, -\varepsilon^{-1}, -1, -\varepsilon, -2.$$

Diese entsprechen den Winkeln

$$(53) \quad \angle(a_1, a_2) = 0^\circ, 36^\circ, 60^\circ, 72^\circ, 90^\circ, 108^\circ, 120^\circ, 144^\circ, 180^\circ.$$

Es kommen also nur Vielfache von 36° , 60° und 90° in Frage. Eine beliebige Vektormenge heißt reduzibel, wenn sie in zwei nicht leere orthogonale Teilmengen zerlegt werden kann, ist das nicht der Fall, dann irreduzibel. Um

eine Einsicht in die Struktur von V_m zu gewinnen, genügt es offenbar, die irreduziblen Bestandteile von V_m zu bestimmen. Wir wollen zunächst feststellen, welche ebenen Vektorkonfigurationen in V_m möglich sind. Die Vektoren $a_1, a_2 \subset V_m$ seien linear unabhängig; sie erzeugen eine gewisse Ebene \mathfrak{E} . Welche Vektoren von V_m außer a_1, a_2 liegen noch in \mathfrak{E} ? Ich betrachte alle diese Vektoren und die Winkel zwischen je zwei benachbarten. Da mit a auch $-a$ in V_m liegt, so treten die Winkel paarweise auf und der Vollwinkel 360° zerfällt nach (53) in gerade Vielfache von $36^\circ, 60^\circ$ und 90° :

$$2i \cdot 36^\circ + 2j \cdot 60^\circ + 2k \cdot 90^\circ = 360^\circ.$$

Aus dieser Gleichung folgt

$$i = 5i_1, j = 3j_1, k = 2k_1$$

und damit

$$i_1 + j_1 + k_1 = 1.$$

Es kann also immer nur eine Sorte von Winkeln vorkommen. Indem wir nötigenfalls a_2 durch $-a_2$ ersetzen, kann bereits $a_1 a_2 \geq 0$ vorausgesetzt werden. Der Fall $a_1 a_2 = \varepsilon^{-1}$ kann durch Ersetzen von a_2 durch $\varepsilon^{-1}(a_1 + a_2) \subset V_m$ auf den Fall $a_1 a_2 = \varepsilon$ zurückgeführt werden. Entsprechend den Werten

$$a_1 a_2 = 0, 1, \varepsilon$$

sind nun genau die drei folgenden Konfigurationen möglich:

1. Ein Viereck für $a_1 a_2 = 0$:

$$(54) \quad \pm a_1, \pm a_2 \text{ (reduzibel).}$$

2. Ein Sechseck für $a_1 a_2 = 1$:

$$(55) \quad \pm a_1, \pm a_2, \pm (a_1 - a_2).$$

3. Ein Zehneck für $a_1 a_2 = \varepsilon$:

$$(56) \quad \pm a_1, \pm a_2, \pm (\varepsilon a_1 - a_2), \pm (a_1 - \varepsilon a_2), \pm \varepsilon (a_1 - a_2).$$

Diese Überlegungen veranlassen uns, allgemein Vektordiagramme W_k zu betrachten, für welche wir fordern:

1. W_k ist irreduzibel und enthält k linear unabhängige Vektoren.

2. Für zwei Vektoren aus W_k kommen andere Werte des Skalarproduktes als 0, $\pm \varepsilon^v$ ($v = -1, 0, 1$), ± 2 nicht in Betracht. Insbesondere ist $a^2 = 2$ für alle $a \subset W_k$.

3. Sind a_1, a_2 linear unabhängige Vektoren aus W_k , so erzeugen sie in W_k , entsprechend ihrem Skalarprodukt, eine Vier-, Sechs- oder Zehneck-konfiguration.

Diejenigen Diagramme, in welchen nur die Skalarprodukte 0, $\pm 1, \pm 2$ vorkommen, sind aus der Theorie der Lieschen Ringe bekannt und sollen hier nicht mehr betrachtet werden. Wir wollen daher noch verlangen:

4. W_k enthält eine Zehneckkonfiguration.

Es wird sich herausstellen, daß es nur zu den Dimensionen $k = 2, 3, 4$ ein Diagramm W_k gibt, und zwar jeweils genau eins. Für $k = 2$ ist diese Behauptung bewiesen. Wir diskutieren den Fall $k = 3$.

Nach 4. liegt in W_3 ein Diagramm $W_2 = [a_1, a_2]$ mit $a_1 a_2 = \varepsilon$. Wir bestimmen, was nach 1. möglich ist, einen von a_1, a_2 unabhängigen Vektor b , der auf W_2 nicht senkrecht steht. Sei dann etwa $b a_1 \neq 0$. Wir können dann b durch einen Vektor b_1 in der Ebene von a_1 und b ersetzen, so daß $b_1 a_1 = \varepsilon^r$ mit $r = 0$ oder 1 gilt. Dabei hat man noch die Freiheit, b_1 durch $\varepsilon^r a_1 - b_1$ zu ersetzen. Wir benutzen sie, um auf alle Fälle $b_1 a_2 \neq 0$ zu erreichen; das soll vorausgesetzt werden; es ist dann $b_1 a_2 = \pm \varepsilon$. Da nun $\angle(a_1, a_2) = 36^\circ$, $\angle(b_1, a_1) \leq 60^\circ$, so ist $\angle(b_1, a_2) \leq 96^\circ$. Es gilt daher $b_1 a_2 = \varepsilon^u$. Für $r = 1$ scheidet $u = -1$ aus, da b_1 nicht in der Ebene von a_1, a_2 liegt. In den fünf übrigen Fällen kann eine Linearkombination a_3 derart bestimmt werden, daß $a_1 a_3 = a_2 a_3 = \varepsilon$ gilt. Entsprechend den Exponentenpaaren $(v, \mu) = (1, 1), (1, 0), (0, 1), (0, 0), (0, -1)$ wähle man $a_3 = b_1, -b_1 + \varepsilon a_1, -b_1 + \varepsilon a_2, \varepsilon^{-1}(a_1 + a_2 - b_1), \varepsilon^{-1}(b_1 + a_2)$. Wir erhalten damit in W_3 Vektoren a_1, a_2, a_3 mit $a_i a_k = \varepsilon$ für $i \neq k$. Jedes Diagramm W_3 , welches diese drei Vektoren enthält, muß nach 3. das System $[a_1, a_2, a_3]$ der Vektoren

$$(57) \quad \pm a_1, \pm(a_1 - \varepsilon a_2), \pm \varepsilon(a_1 - a_2), \pm(a_1 + a_2 - \varepsilon a_3)$$

und aller derjenigen enthalten, die aus (57) durch beliebige Permutation der Ziffern 1, 2, 3 hervorgehen (Gesamtzahl: 30 Vektoren). Die Art und Weise, wie a_3 bestimmt ist, läßt erkennen, daß auch der ursprünglich ausgewählte Vektor b vom Typus eines der Vektoren von (57) ist, also in $[a_1, a_2, a_3]$ enthalten ist. Wir beweisen nun

$$(58) \quad W_3 = [a_1, a_2, a_3],$$

d. h. jeder Vektor aus W_3 kommt unter den Typen (57) vor. Sei b^* ein beliebiger Vektor aus W_3 . b^* steht nicht auf allen a_i senkrecht; sei also etwa $b^* a_1 \neq 0$. Liegt b^* in der Ebene von a_1, a_2 , dann ist $b^* \subset W_2 = [a_1, a_2] \subset [a_1, a_2, a_3]$. Ist b^* dagegen von a_1, a_2 linear unabhängig, dann bestimmt man wie oben a_3^* in W_3 derart, daß $a_1 a_3^* = a_2 a_3^* = \varepsilon$ und $b^* \subset [a_1, a_2, a_3^*]$. Löst man mit dem Ansatz $a_3^* = a_1 a_1 + a_2 a_2 + a_3 a_3$ die Gleichungen

$$a_1 a_3^* = a_2 a_3^* = \varepsilon, a_3^{*2} = 2,$$

so erhält man die Vektoren $a_3^* = a_3$ und $\frac{2\varepsilon}{2+\varepsilon}(a_1 + a_2) - a_3$. Der zweite Vektor hat mit a_3 ein unzulässiges Skalarprodukt, so daß also $a_3^* = a_3$, womit (58) bewiesen ist. Wir kommen jetzt zur Dimension $k = 4$. Auf Grund von 1. und 4. muß in W_4 ein Diagramm $W_3 = [a_1, a_2, a_3]$ mit $a_i a_k = \varepsilon$ für $i \neq k$ enthalten sein. Wir bestimmen einen Vektor b in W_4 , der nicht in W_3

liegt und auf W_3 nicht senkrecht steht. Wir nehmen wieder an, daß $b a_1 \neq 0$. Wie wir gesehen haben, kann dann ein Vektor b_1 mit $a_1 b_1 = a_2 b_1 = \varepsilon$ aufgefunden werden, so daß $b \in [a_1, a_2, b_1]$. Es ist daher auch b_1 nicht in W_3 enthalten, da sonst auch b in W_3 liegen würde. Geometrisch ist nun sofort zu sehen, daß nur noch $a_3 b_1 = \varepsilon$ oder 1 möglich ist. Im zweiten Fall setze man $a_2^* = -b_1 + \varepsilon a_1$, $a_1^* = b_1 - \varepsilon a_2 + \varepsilon a_3$ und findet für diese Vektoren aus W_3 die Skalarprodukte $a_1 a_2^* = a_1 a_1^* = a_3 a_2^* = a_3 a_1^* = a_2^* a_1^* = \varepsilon$. In allen Fällen ist damit in W_4 ein System von vier Vektoren zu ermitteln, in welchem das Skalarprodukt von je zweien gleich ε ist. Für die Vektoren dieses Systems wählen wir nunmehr die Bezeichnung a_1, a_2, a_3, a_4 . Es gilt somit $a_i a_k = \varepsilon$ für $i \neq k$. W_4 muß nach 3. notwendig das System $[a_1, a_2, a_3, a_4]$ der Vektoren

$$(59) \quad \begin{aligned} & \pm a_1, \pm (a_1 - \varepsilon a_2), \pm \varepsilon (a_1 - \varepsilon a_2), \pm (a_1 + a_2 - \varepsilon a_3), \\ & \pm (\varepsilon^{-1} (a_1 + a_2 + a_3) - a_4), \pm (\varepsilon^{-1} (a_1 + a_2 + a_3) - 2 a_4), \\ & \pm (a_1 + a_2 - \varepsilon^{-1} (a_3 + a_4)), \pm (a_1 + a_2 - \varepsilon a_3 - \varepsilon^{-1} a_4) \end{aligned}$$

und aller derjenigen enthalten, die aus den Vektoren (59) durch beliebige Permutation der Ziffern 1, 2, 3, 4 entstehen (Gesamtzahl: 120 Vektoren). Auch jetzt gilt wieder

$$(60) \quad W_4 = [a_1, a_2, a_3, a_4].$$

Hängt ein beliebiger Vektor $b \in W_4$ linear von a_1, a_2, a_3 ab, so liegt wegen der Eindeutigkeit von W_3 der Vektor b in $[a_1, a_2, a_3]$. Wenn dagegen b von den Vektoren a_1, a_2, a_3 linear unabhängig und etwa $b a_1 \neq 0$ ist, dann bestimmt man einen Vektor $b_1 \in W_4$ derart, daß $a_1 b_1 = a_2 b_1 = \varepsilon$ und $b \in [a_1, a_2, b_1]$. Es ist dann notwendig $a_3 b_1 = \varepsilon$ oder 1. Mit dem Ansatz $b_1 = a_1 a_1 + a_2 a_2 + a_3 a_3 + a_4 a_4$, $b_1^2 = 2$ erhält man im ersten Fall als Lösungen von $a_1 b_1 = a_2 b_1 = a_3 b_1 = \varepsilon$ die Vektoren $b_1 = a_4$ und $\varepsilon^{-1} (a_1 + a_2 + a_3) - a_4$, im zweiten Fall als Lösungen von $a_1 b_1 = a_2 b_1 = \varepsilon$, $a_3 b_1 = 1$ die Vektoren $b_1 = \varepsilon^{-1} (a_1 + a_2 + a_4) - a_3$ und $a_1 + a_2 - \varepsilon^{-1} (a_3 + a_4)$. Damit ist erkannt, daß b_1 und folglich auch b in $[a_1, a_2, a_3, a_4]$ liegen, womit (60) bewiesen ist. Die Einsicht, daß das System (60) auch wirklich die Forderungen 1. bis 4. erfüllt, gewinnen wir am einfachsten, wenn wir beachten, daß das Diagramm V_4 zu Ω_4 die Forderungen 1. bis 4. erfüllt und 120 Vektoren besitzt, wie sich später herausstellt. Damit haben wir alle W_k bestimmt; denn W_5 existiert nicht mehr, wie sich aus der folgenden Überlegung ergibt. In jedem (hypothetischen) Diagramm W_5 finden wir zunächst ein Diagramm $W_4 = [a_1, a_2, a_3, a_4]$ mit $a_i a_k = \varepsilon$ für $i \neq k$. Sei b ein Vektor aus W_5 , der nicht in W_4 liegt und auf W_4 nicht senkrecht steht. Nehmen wir etwa wieder $b a_1 \neq 0$ an, so kann wie oben ein Vektor b_1 in W_5 gefunden werden, so daß $a_1 b_1 = a_2 b_1 = \varepsilon$ und $b \in [a_1, a_2, b_1]$ gilt. Es bestehen dann nur die Möglichkeiten

$$a_3 b_1 = \varepsilon \text{ oder } 1, \quad a_4 b_1 = \varepsilon \text{ oder } 1.$$

Der Fall $a_3b_1 = a_4b_1 = \varepsilon$ scheidet aus, weil sonst die Vektoren $a_1 + a_2 + -\varepsilon^{-1}(a_3 + a_4)$ und $a_1 + a_2 - \varepsilon^{-1}(a_3 + b_1)$ mit einem unzulässigen Skalarprodukt in W_5 enthalten wären. Wenn aber $a_3b_1 = a_4b_1 = 1$, dann ist $(a_1 + a_2 - \varepsilon^{-1}(a_3 + a_4))b_1 = 2$, also $b_1 = a_1 + a_2 - \varepsilon^{-1}(a_3 + a_4)$ und daher $b \subset W_4$, was nicht sein sollte. Aus Gründen der Symmetrie brauchen wir nur noch den Fall $a_3b_1 = \varepsilon$, $a_4b_1 = 1$ zu betrachten. Hierfür ist $(\varepsilon^{-1}(a_1 + a_3 + a_4) - a_4)b_1 = 2$, also $b_1 = \varepsilon^{-1}(a_1 + a_2 + a_3) - a_4$, woraus wieder der Widerspruch $b \subset W_4$ folgt.

Wir diskutieren jetzt die Vektordiagramme V_m zu den quadratischen Formen $Q(x_1, \dots, x_m)$ für die Veränderlichenzahlen $m = 4$ und 8. In den Fourierreihen zur Thetareihe von $Q(x_1, \dots, x_m)$ und zur Eisensteinreihe $G_{\frac{m}{2}}(\tau)$ stimmen die Konstanten Koeffizienten überein, daher folgt aus (27)

nach einem wiederholt angewendeten Schluß, daß auch die Koeffizienten zum Exponenten 1 übereinstimmen. Daher ist die Anzahl der Vektoren in V_m gleich 120 bzw. 240. Das einzige Vektordiagramm mit der Vektorenanzahl 120 bzw. 240, in welchem keine Zehneckkonfiguration liegt, ist das in der Theorie der Lieschen Ringe mit A_8 bezeichnete Diagramm⁶⁾. Kommt dagegen eine Zehneckkonfiguration vor, so sind $V_4 = W_4$ und $V_8 = W_4 + W_4$ die einzig möglichen Diagramme. Wenn U_m , $m = 4$ bzw. 8, das Diagramm W_4 bzw. $W_4 + W_4$ oder A_8 enthält, dann bilden bereits m Vektoren aus W_4 bzw. $W_4 + W_4$ oder A_8 eine Fundamentalmasche von U_m vom Volumen 1. Daher ist jede gerade Form $Q(x_1, \dots, x_m)$ mit Determinante ε^{2^r} für $m = 4$ und 8 in einer der Formenklassen Ω_4 , $\Omega_4 + \Omega_4$, Ω_8 enthalten. Die Klassen $\Omega_4 + \Omega_4$ und Ω_8 sind verschieden, weil $W_4 + W_4$ reduzibel und A_8 irreduzibel ist. Wir fassen die Resultate dieses Paragraphen zusammen in

Satz 4. Eine beliebige positiv definite ganzzahlige gerade quadratische Form $Q(x_1, \dots, x_m)$ über $R(\sqrt{5})$ mit Determinante ε^{2^r} von $m = 4$ bzw. 8 Veränderlichen ist in der Formenklasse Ω_4 bzw. einer der Klassen $\Omega_4 + \Omega_4$, Ω_8 enthalten. Die Klassen $\Omega_4 + \Omega_4$ und Ω_8 sind voneinander verschieden; Ω_8 enthält Formen mit ganzrationalen Koeffizienten.

(Eingegangen am 2. 10. 1940.)

Über die Existenz von Algebren beliebigen Ranges mit quadratischer Normenform.

Von

Ernst-August Behrens in Hamburg.

Die mit positiv definiten quadratischen Formen in $n = 2k$ Variablen gebildeten Thetareihen sind Modulfunktionen der Dimension $-k$ und einer Stufe, die mit dem höchsten Elementarteiler der Koeffizientenmatrix der quadratischen Form zusammenhängt. Wie Herr Hecke¹⁾ gezeigt hat, haben diese Thetareihen eine algebraische Eigenschaft, die sich am bequemsten für die den Thetareihen entsprechenden Dirichletreihen formulieren läßt. Nach Adjunktion gewisser kommutativer Matrizen besitzt nämlich das System der Dirichletreihen, deren Thetareihen zu fester Dimension und Stufe gehören, ein Eulerprodukt.

Dies legt die Vermutung nahe, daß diese positiven quadratischen Formen die Normenformen zu gewissen hyperkomplexen Systemen sind (oder mit solchen zusammenhängen), wie das bekanntlich bei 2 und 4 Variablen tatsächlich der Fall ist. In der vorliegenden Arbeit soll untersucht werden, ob Systeme mit derartigen Normenformen existieren.

Bei den quadratischen Zahlkörpern und den Quaternionen sind die Normen selbst binäre bzw. quaternäre quadratische Formen. Sie sind folgendermaßen gebildet: Diese Algebren besitzen einen Antiautomorphismus derart, daß das Produkt des allgemeinen Elements mit seinem Bilde eine quadratische Form, multipliziert mit dem Einselement der Algebra, ist. Die Koeffizienten der Form liegen im Grundkörper P der Algebra. Diese Normenform ist multiplikativ, d. h. das Produkt der Normen zweier Elemente ist gleich der Norm ihres Produktes.

Auch für Algebren beliebigen Ranges n hat man eine Norm definiert. $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ seien die Basiselemente der Algebra mit den Strukturkonstanten c_{ik}^l ,

$$(1) \quad \varepsilon_i \varepsilon_k = \sum_{l=1}^n c_{ik}^l \varepsilon_l,$$

$c_{ik}^l \in P$, P Grundkörper der Charakteristik 0 der Algebra.

¹⁾ E. Hecke, Über Modulfunktionen und die Dirichletreihen mit Eulerscher Produktentwicklung I und II. Math. Annalen 114 (1937).

Als Norm des allgemeinen Elementes $a = \sum_{k=1}^n x_k \varepsilon_k$ definiert man die Determinante derjenigen Matrix, die bei der regulären Darstellung D der Algebra dem Element a zugeordnet ist:

$$N_D(a) = \left| \sum_k c'_{ik} x_k \right|.$$

Für $n = 4$ ist diese „Darstellungsnorm“ $N_D(a)$ das Quadrat der gewöhnlichen Quaternionennorm $N(a)$. Im allgemeinen ist aber $N_D(a)$ weder eine quadratische Form, wie bei $n = 2$, noch Potenz irgendeiner quadratischen Form, wie es bei $n = 4$ der Fall ist. Das ist der Inhalt von Satz 3.

Will man also für Algebren beliebigen Ranges n als Norm eine quadratische Form $f(x_1, \dots, x_n)$ definieren, muß man offenbar diese Eigenschaft

$$N_D(a) = (f(x_1, \dots, x_n))^2$$

fallen lassen.

Aber auch auf die beiden anderen obengenannten Eigenschaften der Norm, die sie im binären und quaternären Fall besitzt, muß man dann verzichten. Satz 1 zeigt nämlich, daß es für beliebigen Rang $n \neq 2$ oder 4 nicht nur keinen Antiautomorphismus, sondern nicht einmal eine lineare Abbildung der Algebra in sich gibt mit der Eigenschaft, daß das Produkt des allgemeinen Elementes mit seinem Bilde eine quadratische Form, multipliziert mit dem Einselement der Algebra, ist. Und Satz 2 besagt, daß eine Norm, die durch eine quadratische Form definiert ist, nicht multiplikativ sein kann, wenn $n \neq 2$ oder 4.

Satz 2 folgt für $n > 8$ bereits aus einem Satz von A. Hurwitz²⁾. Da nämlich die Koeffizienten des Produktes zweier Elemente bilinear in den Koeffizienten der beiden Faktoren sind, würde die Multiplikativität der quadratischen Form $f(x_1, \dots, x_n)$ nach sich ziehen das Bestehen einer bilinearen Substitution

$$z_i = \sum_{k=1}^n c'_{ik} x_i y_k$$

mit der Eigenschaft

$$f(x_1, \dots, x_n) \cdot f(y_1, \dots, y_n) = f(z_1, \dots, z_n).$$

Hurwitz zeigt, daß dies außer in den Fällen $n = 2, 4$ oder 8 unmöglich ist. Zum Beweis des Satzes 2 brauchen wir dann nur noch den Fall $n = 8$ zu berücksichtigen.

Das Prinzip ist bei den Beweisen aller unserer Sätze das gleiche. Da wir aus arithmetischen Gründen fordern müssen, daß die „Norm“ der Eins nicht

²⁾ A. Hurwitz, Über die Composition der quadratischen Formen von beliebig vielen Variablen. Gött. Nachr. 1898, S. 309.

verschwinden soll, können wir nach der Transformationstheorie quadratischer Formen annehmen, daß f Diagonalgestalt hat,

$$f(x_1, \dots, x_n) = \sum_{i=1}^n d_i x_i^2,$$

und daß das erste Basiselement ε_1 gleich Eins ist,

$$(2) \quad \varepsilon_1 \varepsilon_k \quad \varepsilon_k \varepsilon_1 = \varepsilon_k.$$

In jedem der Sätze führt dann die für $f(x_1, \dots, x_n)$ geforderte Eigenschaft auf eine Algebra mit den Multiplikationsregeln

$$(2) \quad \varepsilon_1 \varepsilon_k = \varepsilon_k \varepsilon_1 = \varepsilon_k,$$

$$(8) \quad \varepsilon_i \varepsilon_k = -\varepsilon_k \varepsilon_i, \text{ wenn } i \neq k \text{ und beide ungleich } 1,$$

$$(9) \quad \varepsilon_i^2 = -\varrho_i \varepsilon_1, \quad \varrho_i \in \mathbb{P}, \text{ wenn } i \neq 1.$$

Mit Hilfe der Darstellungstheorie zeigen wir dann, daß diese Algebra nicht assoziativ wäre außer für $n = 2$ oder 4 .

Der Darstellungstheorie bedienen sich auch P. Jordan, J. v. Neumann und E. Wigner³⁾ bei ihrem Beweis des Satzes von Hurwitz. Sie kommen auf das gleiche System der ε_i wie wir und fragen nach gewissen reellen Darstellungen. Ihre Überlegungen sind aber nicht ganz einwandfrei, weil ihnen die Annahme zugrunde liegt, daß die ε_i mit diesen Multiplikationsregeln eine Gruppe erzeugen. Die Verfasser übersehen dabei, daß die Gruppe nicht assoziativ wäre. Dies Versehen hat aber auf die Ergebnisse ihrer Arbeit keinen Einfluß.

Satz 1. *Außer für $n = 1, 2$ und 4 gibt es keine Algebra mit einer linearen Abbildung in sich derart, daß das Produkt des allgemeinen Elements mit seinem Bilde eine quadratische Form multipliziert mit dem Einselement ε_1 der Algebra ist.*

Beweis. Bei der linearen Abbildung werde das Element $a = \sum_i x_i \varepsilon_i$ abgebildet auf

$$(3) \quad \bar{a} = \sum_{k,r} m_{kr} x_r \varepsilon_k, \quad m_{kr} \in \mathbb{P}.$$

Angenommen, es wäre

$$\begin{aligned} a \bar{a} &= \sum_{i,k,r} m_{kr} x_i x_r \varepsilon_i \varepsilon_k = \sum_{i,k,r,l} c_{ik}^l m_{kr} x_i x_r \varepsilon_l \\ &= \varepsilon_1 \cdot f(x_1, \dots, x_n) = \sum_i d_i x_i^2 \varepsilon_1. \end{aligned}$$

³⁾ On an algebraic generalisation of the quantum mechanical formalism. *Annals of Mathematics* 35 (1934), S. 51.

Aus dem Vergleich der Koeffizienten der Basiselemente folgt

$$\sum_{i,k,r} c_{ik}^i m_{kr} x_i x_r = \delta_{11} \sum_i d_i x_i^2$$

mit $\delta_{ik} = \begin{cases} 1, & \text{wenn } i = k \\ 0, & \text{wenn } i \neq k, \end{cases}$

und hieraus wieder durch Koeffizientenvergleich der Unbestimmten x_i

$$(4) \quad \sum_k (c_{ik}^i m_{kr} + c_{rk}^i m_{ki}) = 2 \delta_{11} \delta_{ir} d_i.$$

Da $\varepsilon_1 \varepsilon_k = \varepsilon_k \varepsilon_1 = \varepsilon_1$ ist, können wir einige Strukturkonstanten ohne weiteres angeben:

$$(5) \quad c_{1k}^i = c_{k1}^i = \delta_{ki}.$$

Diese können wir dann zur Berechnung der Koeffizienten m_{kr} der linearen Abbildung (3) benutzen. Wir setzen zunächst in (4) $i = r = 1$:

$$\sum_k c_{1k}^1 m_{k1} = m_{11} = \delta_{11} d_1,$$

und dann i beliebig und $r = 1$:

$$\begin{aligned} \sum_k (c_{ik}^i m_{k1} + c_{1k}^i m_{ki}) &= 2 \delta_{11} \delta_{i1} d_i \\ m_{1i} &= 2 \delta_{11} \delta_{i1} d_i - \delta_{i1} d_1. \end{aligned}$$

Also zusammengefaßt:

$$(6) \quad \begin{aligned} m_{11} &= d_1, \\ m_{ii} &= -d_1, \text{ wenn } i \neq 1, \\ m_{1i} &= 0, \text{ wenn } i \neq 1, \end{aligned}$$

oder

$$\bar{a} = d_1(x_1 \varepsilon_1 - x_2 \varepsilon_2 - \dots - x_n \varepsilon_n).$$

Es sei noch bemerkt, daß für $N(\varepsilon_1) = 1$ nach der Transformationstheorie $d_1 = 1$ wird.

Weil alle m_{kr} bekannt sind, können wir die Beziehung (4) zwischen den Strukturkonstanten vereinfachen zu:

$$(7) \quad c_{ir}^i m_{rr} + c_{ri}^i m_{ii} = 2 \delta_{11} \delta_{ir} d_i.$$

Sind nun r und i voneinander verschieden und beide ungleich 1, so erhalten wir wegen $m_{rr} = m_{ii}$

$$c_{ir}^i = -c_{ri}^i$$

und daraus nach (1) für die Basiselemente

$$(8) \quad \varepsilon_i \varepsilon_r = -\varepsilon_r \varepsilon_i, \quad \text{wenn } i \neq r \text{ und } i, r \neq 1.$$

Ist aber $i = r \neq 1$, dann ergibt sich aus (7)

$$c_{ii}^i = -\delta_{11} \frac{d_i}{d_1}$$

und damit eine Aussage über die Quadrate der von ε_1 verschiedenen Basis-elemente

$$(9) \quad \varepsilon_i^2 = -\frac{d_i}{d_1} \varepsilon_1, \quad \text{wenn } i \neq 1.$$

Satz 1 folgt dann sofort aus dem

Hilfssatz: Nur für $n = 2$ und 4 gibt es Algebren mit den Eigenschaften

$$(2) \quad \varepsilon_1 \varepsilon_k = \varepsilon_k \varepsilon_1 = \varepsilon_k,$$

$$(8) \quad \varepsilon_i \varepsilon_r = -\varepsilon_r \varepsilon_i, \quad \text{wenn } i \neq r \text{ und } i, r \neq 1,$$

$$(9) \quad \varepsilon_i^2 = -\frac{d_i}{d_1} \varepsilon_1, \quad \text{wenn } i \neq 1.$$

Beweis. Für einen anderen Rang n wäre die Algebra nicht assoziativ. Nämlich ausgedrückt in den Strukturkonstanten der Algebra lautet das Assoziativgesetz

$$(10) \quad \sum_l c_{ik}^l c_{ul}^r = \sum_m c_{um}^r c_{mk}^i.$$

Eine bequeme Übersicht über die Strukturkonstanten können wir uns durch die reguläre Darstellung der Algebra

$$\begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix} \cdot \alpha = (a_{i1}) \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

verschaffen. Die den Basiselementen $\alpha = \varepsilon_k$ zugeordneten Matrizen (a_{i1}) sind gerade die zu Matrizen M_k vereinigten Strukturkonstanten c_{ik}^l

$$\begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix} \cdot \varepsilon_k = (c_{ik}^l) \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix} = M_k \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

Gemäß den Darstellungseigenschaften gehen die Gleichungen (8) und (9) über in

$$(11) \quad M_i M_r = -M_r M_i, \quad \text{wenn } i \neq r \text{ und } i, r \neq 1,$$

$$(12) \quad M_i^2 = -\frac{d_i}{d_1} E, \quad \text{wenn } i \neq 1.$$

Aus (12) folgt für die Determinante

$$(13) \quad |M_i| \neq 0,$$

und aus (11): $|M_i| |M_r| = (-1)^n |M_r| |M_i|$. Also muß der Rang n der Algebra gerade sein. Daß er auch nicht von 2 und 4 verschieden sein kann, beweisen wir mit Hilfe des Assoziativgesetzes (10). Aus ihm folgt nämlich, daß in allen anderen Fällen die Determinante $|M_i M_r| = 0$ ist, was (13) widerspricht.

Um hierzu (10) in Matrizen schreiben zu können, müssen wir c_{ui}^v durch c_{iu}^v und c_{ui}^u durch c_{iu}^u ausdrücken. Es sei $u \neq 1$, ebenso i und v . Dann bekommen wir für die linke Seite von (10) unter Benutzung der sich aus (2), (8) und (9) für die Strukturkonstanten ergebenden Relationen

$$\begin{aligned}\sum_i c_{ik}^j c_{ui}^v &= - \sum_{i \neq 1} c_{ik}^j c_{iu}^v + c_{ik}^1 c_{u1}^v \\ &= - \sum_{i=1}^n c_{ik}^j c_{iu}^v + 2 \delta_{uv} c_{ik}^1;\end{aligned}$$

und ebenso folgt für die rechte Seite:

$$\sum_m c_{ui}^m c_{mk}^v = - \sum_m c_{iu}^m c_{mk}^v + 2 \delta_{kv} \delta_{iu} c_{ii}^1;$$

also zusammen:

$$\begin{aligned}- \sum_i c_{ik}^j c_{iu}^v + 2 \delta_{uv} c_{ik}^1 &= - \sum_m c_{iu}^m c_{mk}^v + 2 \delta_{kv} \delta_{iu} c_{ii}^1, \\ (14) \quad \sum_m c_{iu}^m c_{mk}^v - \sum_i c_{ik}^j c_{iu}^v &= 2 \delta_{kv} \delta_{iu} c_{ii}^1 - 2 \delta_{uv} c_{ik}^1, \quad \text{wenn } i, v, u \neq 1.\end{aligned}$$

Auf der linken Seite von (14) steht dasjenige Element a_{iv} , das in der i -ten Zeile und v -ten Spalte der Matrix

$$(15) \quad M_u M_k - M_k M_u = (a_{iv})$$

steht. Abgesehen von der ersten Zeile und der ersten Spalte sind also die Matrizenelemente

$$(16) \quad a_{iv} = 2 \delta_{kv} \delta_{iu} c_{ii}^1 - 2 \delta_{uv} c_{ik}^1.$$

Auch für i oder $v = 1$ läßt sich ein ähnlicher Ausdruck angeben; wir müssen dann mehrere Fälle unterscheiden. Da die Kenntnis der Elemente der ersten Zeile und ersten Spalte für das Folgende nicht nötig ist, kann die Berechnung hier übergangen werden.

Wir nehmen $u, k \neq 1$ und $u \neq k$. Wegen (11) ist dann die linke Seite von (15)

$$M_u M_k - M_k M_u = 2 M_u M_k$$

und somit

$$(17) \quad M_u M_k = \left(\frac{a_{iv}}{2}\right).$$

Für $n > 4$ gibt es nun in dieser Matrix zwei linear abhängige Spalten. Es ist dann nämlich möglich, für v zwei Zahlen v_1 und v_2 zu wählen, die von 1, k und u verschieden sind. Nach (16) stehen in den Spalten v_1 und v_2 höchstens in der ersten Zeile von Null verschiedene Elemente. Diese Spalten sind also linear abhängig und damit die Determinante $|M_u M_k| = 0$, was (13) widerspricht. Der Hilfssatz und damit Satz 1 sind also bewiesen.

Satz 2. *Außer für $n = 1, 2$ und 4 gibt es keine Algebra mit multiplikativer quadratischer Normenform.*

Beweis. Wieder seien $a = \sum_i x_i \varepsilon_i$ und $b = \sum_k y_k \varepsilon_k$ Elemente der Algebra, $\varepsilon_1 = 1$ und $N(a) = \sum_i d_i x_i^2$;

$$(18) \quad N(ab) = N(a)N(b).$$

Die Norm des Produktes der Elemente

$$ab = \sum_{i,k} x_i y_k \varepsilon_i \varepsilon_k = \sum_{i,k,l} c_{ik}^l x_i y_k \varepsilon_l$$

ist:

$$\begin{aligned} N(ab) &= \sum_l d_l \left(\sum_{i,k} c_{ik}^l x_i y_k \right)^2 \\ &= \sum_l d_l \left(\sum_{i,k} c_{ik}^l x_i y_k \right) \left(\sum_{j,u} c_{ju}^l x_j y_u \right) \\ &= \sum_{i,k} \sum_l d_l (c_{ik}^l)^2 x_i^2 y_k^2 + \sum_{i,j} \sum_{k,u} d_l c_{ik}^l c_{ju}^l x_i x_j y_k y_u, \end{aligned}$$

und das Produkt der Normen ist:

$$N(a)N(b) = \sum_{i,k} d_i d_k x_i^2 y_k^2.$$

Aus (18) folgt dann durch Vergleich der Koeffizienten linear unabhängiger $x_i x_j y_k y_u$

$$(19) \quad \sum_l d_l (c_{ik}^l c_{ju}^l + c_{iu}^l c_{jk}^l) = 2 \delta_{ij} \delta_{ku} d_i d_k.$$

Aus dieser letzten Gleichung folgen nun die Relationen (8) und (9) zwischen den Basiselementen der Algebra und so nach dem Hilfssatz die Behauptung von Satz 2.

Wir erweitern dazu den Grundkörper \mathbf{P} der Algebra durch Adjunktion der Quadratwurzeln aus allen d_i . Als Basiselemente dieser neuen Algebra wählen wir

$$\eta_i = \frac{\varepsilon_i}{\sqrt{d_i}}.$$

Die neuen Multiplikationskonstanten ergeben sich dann aus den alten durch

$$\gamma_{ik}^l = c_{ik}^l \sqrt{\frac{d_l}{d_i d_k}},$$

und Gleichung (19) wird äquivalent mit

$$\sum_l (\gamma_{ik}^l \gamma_{ju}^l + \gamma_{iu}^l \gamma_{jk}^l) = 2 \delta_{ij} \delta_{ku}.$$

Wieder vereinigen wir die Multiplikationskonstanten γ_{ik}^l zu den Bildern M_k der Basiselemente η_k bei der regulären Darstellung. Für (19) können wir dann in diesen Matrizen schreiben

$$(20) \quad M_k M'_u = -M_u M'_k, \text{ wenn } k \neq u,$$

$$(21) \quad M_k M'_k = E.$$

Wegen $\eta_1 = \frac{e_1}{\sqrt{d_1}}$ ist $M_1 = \frac{1}{\sqrt{d_1}} E$; dies zieht, wenn wir in (20) $k = 1$ setzen, die Schiefsymmetrie der Matrizen $M'_u = -M_u$ für $u \neq 1$ nach sich. Also

$$M_u M_k = -M_k M_u, \text{ wenn } k \neq u \text{ und } k, u \neq 1,$$

$$M_k^2 = E.$$

Da die M_k die Bilder der Basiselemente η_k bei einer isomorphen Darstellung sind, ist

$$\eta_1 \eta_k = \eta_k \eta_1 = \frac{\eta_k}{\sqrt{d_1}},$$

$$\eta_u \eta_k = -\eta_k \eta_u, \text{ wenn } k \neq u \text{ und } k, u \neq 1,$$

$$\eta_k^2 = \eta_1 \sqrt{d_1}.$$

Schließlich führen wir die neuen Basiselemente $\zeta_1 = \eta_1 \sqrt{d_1} = e_1$, $\zeta_2 = \eta_2$, ..., $\zeta_n = e_n$ ein und erhalten ein System mit

$$(2) \quad \zeta_1 \zeta_k = \zeta_k \zeta_1 = \zeta_k,$$

$$(8) \quad \zeta_u \zeta_k = -\zeta_k \zeta_u, \text{ wenn } k \neq u \text{ und } k, u \neq 1,$$

$$(9) \quad \zeta_k^2 = \zeta_1.$$

Die Nichtassoziativität dieses Systems folgt dann sofort aus dem Hilfssatz.

Unter der Darstellungsnorm $N_D(a)$ verstehen wir die Determinante derjenigen Matrix $(\sum_{k=1}^n c_{ik}^j x_k)$, die das Bild des Elementes $a = \sum_{k=1}^n x_k e_k$ bei der regulären Darstellung der Algebra ist.

Satz 3. *Außer für $n = 1, 2$ und 4 gibt es keine Algebra vom Range n , deren Darstellungsnorm $N_D(a)$ Potenz einer irgendwie als Normenform $N(a)$ definierten quadratischen Form $f(x_1, \dots, x_n)$ ist.*

Beweis. Es sei $N_D(a) = N^m(a) = (f(x_1, \dots, x_n))^m$. Da $N_D(a)$ eine homogene Form in x_1, \dots, x_n vom Grade n ist, muß n -gerade und $n = 2m$ sein. Aus der Multiplikativität der Darstellungsnorm,

$$N_D(a \cdot b) = N_D(a) N_D(b),$$

folgt

$$N^m(a \cdot b) = N^m(a) N^m(b),$$

also

$$N(a \cdot b) = N(a) N(b) \cdot \zeta_m,$$

wobei ζ_m eine m -te Einheitswurzel ist. Nötigenfalls erweitern wir dazu den Grundkörper P der Algebra zum Körper der m -ten Einheitswurzeln über P .

Zu jeder Kombination a, b gibt es also eine m -te Einheitswurzel

$$\zeta_m = \zeta_m(a, b) = \frac{N(a, b)}{N(a)N(b)}.$$

Die rechte Seite dieser Gleichung können wir als rationale Funktion in den $4m$ Variablen $x_1, \dots, x_{2m}, y_1, \dots, y_{2m}$ auffassen:

$$\zeta_m(a, b) = \frac{f\left(\sum_{i,k} c_{ik}^1 x_i y_k, \dots, \sum_{i,k} c_{ik}^m x_i y_k\right)}{f(x_1, \dots, x_{2m}) \cdot f(y_1, \dots, y_{2m})},$$

$$f(x_1, \dots, x_{2m}) = \sum_i d_i x_i^2.$$

Wenn wir für die Variablen Elemente aus P einsetzen, nimmt diese rationale Funktion höchstens m verschiedene Werte an, nämlich die m -ten Einheitswurzeln. Setzen wir voraus, daß die Charakteristik von P ungleich 0 ist, muß diese Funktion eine Konstante, also $\zeta_m(a, b)$ unabhängig von a und b sein. Wir berechnen diese Konstante, wenn wieder die Algebra auf die Gestalt $\varepsilon_1 = 1$ und die Form N auf Diagonalgestalt $\sum_i d_i x_i^2$ gebracht ist, sehr einfach:

$$\zeta_m(a, b) = \zeta_m(\varepsilon_1, \varepsilon_1) = N^{-1}(\varepsilon_1) = d_1^{-1}.$$

$$N(ab) = d_1^{-1} N(a)N(b).$$

Wie beim Beweis von Satz 2 werden wir dann durch Koeffizientenvergleich auf das nichtassoziative System (2), (8), (9) geführt.

(Eingegangen am 23. 9. 1940.)

Die 28 Doppeltangenten einer Kurve vierter Ordnung*).

Von

Li En-Po in Ho-Pei (China).

Einleitung.

Steiner¹⁾ hat eine Reihe von Sätzen über die Doppeltangenten der ebenen Kurven vierter Ordnung ohne Beweise ausgesprochen. Einen Teil dieser Sätze haben Hesse, Salmon und Noether bewiesen. Diese Arbeit stellt sich das Ziel, alle von Steiner ausgesprochenen Sätze zu beweisen und noch einige weitergehende Ergebnisse herzuleiten. Die betrachtete Kurve vierter Ordnung wird dabei als doppelpunktfrei und irreduzibel vorausgesetzt.

Wie dies zuerst Hesse²⁾ getan hat, betrachten wir die Koordinaten eines Punktes in der Ebene als die Parameter eines Netzes von quadratischen Flächen im Raum und das Verschwinden der Diskriminante als die Gleichung der Kurve vierter Ordnung. Dieses Verschwinden ist die Bedingung dafür, daß die Fläche des Netzes ein Kegel ist; es entspricht also jedem Punkt der Kurve vierter Ordnung in der Ebene ein Kegel des Netzes im Raum. Hieraus hat Hesse viele Ergebnisse erhalten und für die Untersuchung der Doppeltangenten verwertet. Wir können aber auch ohne Betrachtung der Kegel, nur mit Hilfe der acht Schnittpunkte der drei Basisflächen und der invarianten Eigenschaften der Polaren, alle 28 Doppeltangenten darstellen und ihre Eigenschaften diskutieren.

Bekannt ist³⁾, daß die 56 Berührungspunkte der 28 Doppeltangenten auf einer Kurve 14. Ordnung liegen. Wenn man also die Schnittpunkte dieser Kurve mit der Kurve vierter Ordnung bestimmt, kann man alle Doppeltangenten erhalten. Diese Methode ist zwar nicht geeignet zur Erforschung der Eigenschaften der Doppeltangenten, aber ihre Existenz kann man so

*) Diese Arbeit wurde von der philosophischen Fakultät der Universität Leipzig als Dissertation angenommen (D 15).

¹⁾ Steiner, Eigenschaften der Kurven vierten Grades rücksichtlich ihrer Doppeltangenten. Crelles Journal, Bd. 49.

²⁾ Hesse, Über die Doppeltangenten der Kurven vierter Ordnung. Crelles Journal, Bd. 49. — Zu den Doppeltangenten der Kurven vierter Ordnung. Crelles Journal, Bd. 55.

³⁾ Hesse, Transformation der Gleichung der Kurven vierzehnten Grades, welche eine gegebene Kurve vierten Grades in den Berührungspunkten ihrer Doppeltangenten schneiden. Crelles Journal, Bd. 52; vgl. auch Weber, Algebra II, Abschn. 13.

beweisen. (Die Existenz und Anzahl der Doppeltangenten folgen übrigens auch aus den Plückerschen Formeln.)

Salmon⁴⁾ faßte die Kurve vierter Ordnung als die Einhüllende eines Systems $\lambda^2 R + 2\lambda Q + S = 0$ auf, wobei R , Q und S drei Kegelschnitte sind. Dabei zeigte er, daß jedes Doppeltangentenpaar fünf andere Paare bestimmt; die sechs Paare zusammen nennt man eine Steinersche Gruppe. Er zeigt noch, wie man aus den Gleichungen von je drei Paaren aus einer Steinerschen Gruppe die Gleichung der Kurve gewinnen kann. Auf Grund dieser Bemerkung kann man aber weiter zeigen, daß irgendeine Kurve vierter Ordnung als die Diskriminante eines quadratischen Flächennetzes dargestellt werden kann. So kann man von der Salmonschen Erzeugung zur Hesseschen übergehen und umgekehrt. Wir werden das nachher unter 1. 2 näher ausführen.

Steiner hat die Behauptung aufgestellt, daß die 63 Steinerschen Gruppen sich für $2 < n < 8$ nach einem Gesetz zu n und n zu Systemen $S[n]$ ordnen, so daß zu je $(n - 1)$ Gruppen allemal eine und nur eine bestimmte n -te Gruppe gehört; wählt man aus jeder der n Gruppen eines Systems ein beliebiges Paar, so liegen die $4n$ Berührungspunkte auf einer Kurve n -ter Ordnung. Zum Beweise dieser Behauptung kann man ein Kriterium von M. Noether⁵⁾ heranziehen, wodurch man entscheiden kann, ob die $4n$ Berührungspunkte von $2n$ Doppeltangenten von einer Kurve n -ter Ordnung ausgeschnitten werden. Darüber hinaus kann man nachweisen, daß auf dieser Kurve noch $n(n - 2)$ Berührungspunkte einer Kurve $(2n - 4)$ -ter Ordnung mit denselben $2n$ Doppeltangenten liegen.

Um die Steinerschen Systeme von je n Steinerschen Gruppen zu untersuchen, führen wir den Begriff der Punktgruppe auf einer Kurve ein und erklären die Multiplikation von zwei Gruppen. Aus dieser Betrachtung finden wir sehr leicht die Beziehungen zwischen den Steinerschen Gruppen, sowie den Hauptsatz hinsichtlich der Eigenschaften der Systeme $S[n]$. Demnach können wir alle Systeme $S[n]$ ermitteln und die Anzahl der Systeme berechnen. Als Beispiel schreiben wir uns die Systeme für $n = 3$ und $n = 4$ vollständig hin und untersuchen auch die Kurve dritter Ordnung, welche durch die 12 Berührungspunkte der sechs Doppeltangenten eines Systems $S[3]$ geht.

Im Kapitel I werden wir die Ergebnisse von Salmon und Hesse wiedergeben, soweit sie zu unserem Problem Beziehung haben. Kapitel II bringt die Theorie der Punktgruppen und die Beweise der Steinerschen Sätze.

⁴⁾ Salmon-Fiedler, Analytische Geometrie der höheren ebenen Kurven.

⁵⁾ M. Noether, Über die Gleichungen achten Grades und ihr Auftreten in der Theorie der Kurven vierter Ordnung. Math. Annalen, Bd. 15.

I. Hessesche Erzeugung und Steinersche Gruppen.

1. 1. Es sei $\psi \quad \lambda_0 A + \lambda_1 B + \lambda_2 C = 0$ ein Netz von quadratischen Flächen im Raum S_3 ; die Koeffizienten dieser Fläche sind Linearformen in λ :

$$(1) \quad f_{kl} = \lambda_0 a_{kl} + \lambda_1 b_{kl} + \lambda_2 c_{kl}.$$

Faßt man nun die $\lambda_0, \lambda_1, \lambda_2$ als Koordinaten in einer Ebene auf, der Bildebene des Netzes, so definiert die Gleichung $f_{kl} = 0$ in der Bildebene eine Gerade und die Diskriminante

$$(2) \quad F = \begin{vmatrix} f_{11} & f_{12} & f_{13} & f_{14} \\ f_{21} & f_{22} & f_{23} & f_{24} \\ f_{31} & f_{32} & f_{33} & f_{34} \\ f_{41} & f_{42} & f_{43} & f_{44} \end{vmatrix} = 0$$

eine Kurve 4. Ordnung. Jedem Punkt dieser Kurve entspricht ein Kegel des Netzes.

Es ist bekannt, daß drei Quadriken A, B, C im allgemeinen acht verschiedene Schnittpunkte haben, das Netz ψ also acht Basispunkte hat. Eine Ausnahme kann nur dann auftreten, wenn die Quadriken A, B, C eine Kurve (Gerade, Kegelschnitt oder kubische Raumkurve) oder eine Ebene gemeinsam haben oder wenn mehrfache Schnittpunkte vorkommen. In allen diesen Fällen hätten die drei Quadriken in einem ihrer gemeinsamen Punkte eine gemeinsame Tangente. Es gäbe dann eine Linearkombination $D = \lambda_0 A + \lambda_1 B + \lambda_2 C$, die in diesem Punkte einen Doppelpunkt hat, also einen Kegel. Wir werden sehen, daß dieser Fall nicht vorkommen kann, wenn die Bildkurve (2) doppelpunktfrei ausfallen soll.

Lemma. Ein Büschel von Quadriken in S_3 enthält im allgemeinen 4 Kegel. Hat einer der Kegel seinen Doppelpunkt in einem Basispunkt des Büschels, so fallen zwei der vier Kegel in ihm zusammen.

Beweis. Das Büschel sei durch

$$\lambda_1 B + \lambda_2 C = 0$$

gegeben. Die Kegel im Büschel entsprechen dann den Wurzeln der biquadratischen Gleichung

$$(3) \quad \begin{vmatrix} \lambda_1 b_{11} + \lambda_2 c_{11} & \lambda_1 b_{12} + \lambda_2 c_{12} & \dots & \lambda_1 b_{14} + \lambda_2 c_{14} \\ \lambda_1 b_{21} + \lambda_2 c_{21} & \lambda_1 b_{22} + \lambda_2 c_{22} & \dots & \lambda_1 b_{24} + \lambda_2 c_{24} \\ \dots & \dots & \dots & \dots \\ \lambda_1 b_{41} + \lambda_2 c_{41} & \lambda_1 b_{42} + \lambda_2 c_{42} & \dots & \lambda_1 b_{44} + \lambda_2 c_{44} \end{vmatrix} = 0.$$

Ist der Punkt $(1, 0, 0)$ die Spitze des Kegels B , so ist $b_{11} = b_{12} = b_{13} = b_{14} = 0$. Geht nun außerdem die Fläche C durch denselben Punkt, so wird $c_{11} = 0$, und die linke Seite von (3) wird durch λ_2^2 teilbar.

Einem Büschel des Netzes ψ entspricht eine Gerade in der Bildebene des Netzes. Den vier Kegeln des Büschels entsprechen die vier Schnittpunkte der Geraden mit der Kurve (2). Würde das Netz einen Kegel D enthalten, dessen Spitze ein Basispunkt des Netzes wäre, so würde jedes D enthaltende Büschel in D zwei zusammenfallende Kegel haben, also wäre der D entsprechende Punkt in der Bildebene ein Doppelpunkt der Kurve (2). Das ist unmöglich, da die Kurve als doppeltpunktfrei vorausgesetzt wurde. Auf Grund der oben genannten Bemerkung folgt nun:

Das Netz ψ hat 8 verschiedene Basispunkte⁶⁾.

Wählt man irgend vier Basispunkte E_1, E_2, E_3, E_4 des Netzes als Ecken des Koordinatentetraeders, so nimmt die Gleichung (2) die Gestalt

$$(3) F = \begin{vmatrix} 0 & f_{12} & f_{13} & f_{14} \\ f_{21} & 0 & f_{23} & f_{24} \\ f_{31} & f_{32} & 0 & f_{34} \\ f_{41} & f_{42} & f_{43} & 0 \end{vmatrix} = (f_{14}f_{23} - f_{12}f_{34} - f_{13}f_{24})^2 - 4f_{12}f_{34}f_{13}f_{24} = 0$$

an.

Aus (3) sieht man, daß jede der vier Geraden $f_{12} = 0, f_{34} = 0, f_{13} = 0, f_{24} = 0$ die Kurve F zweimal in zwei zusammenfallenden Punkten schneidet und deshalb eine Doppeltangente ist.

Mit $(E: \psi)$ bezeichnet man die erste Polarform irgendeines Punktes E in bezug auf ψ . Dann gilt:

$$(E_1: \psi) = x_2f_{12} + x_3f_{13} + x_4f_{14}$$

und

$$E_2: (E_1: \psi) = f_{12},$$

wobei $E_1 = (1, 0, 0, 0)$ und $E_2 = (0, 1, 0, 0)$.

Jede Polarform der Form ψ in bezug auf je zwei Basispunkte E_i und E_k bestimmt also eine Doppeltangente; die zwei Basispunkte werden die Pole der Doppeltangente genannt. Im ganzen gibt es 28 Doppeltangenten; jede von ihnen bezeichnet man mit $\bar{j}k$, und zwar ist $\bar{j}k = \overline{kj}$.

1.2. Aus dem Obigen folgt, daß die Diskriminante eines Netzes von quadratischen Flächen im Raum S_3 eine Kurve vierter Ordnung darstellt. Nun ergibt sich die Frage: Wie stellt man eine gegebene Kurve vierter Ordnung $F = 0$ als die Diskriminante eines solchen Netzes dar?

Seien t_1, t'_1 irgend zwei verschiedene Doppeltangenten der Kurve F und Q ein beliebiger durch ihre vier Berührungspunkte gehender Kegelschnitt. Dann definiert $\lambda F + Q^2 = 0$ ein Büschel von Kurven vierter Ordnung. Jede Kurve im Büschel berührt die Geraden t_1 und t'_1 in denselben vier Punkten.

⁶⁾ Dieser Beweis wurde einem unpublizierten Manuskript von Prof. van der Waerden entnommen.

Wählt man also λ so, daß die Kurve $\lambda F + Q^2 = 0$ durch den Schnittpunkt der beiden Geraden geht, so hat sie mit jeder der beiden Geraden 5 Punkte gemeinsam, also sind die beiden Geraden als Bestandteile in ihr enthalten. Man hat daher die Identität

$$\lambda F + Q^2 = t_1 t'_1 S$$

oder, wenn man λF wieder F nennt und $t_1 t'_1 = R$ setzt,

$$(4) \quad F = RS - Q^2,$$

wo R , S und Q quadratische Formen sind und R in t_1 und t'_1 zerfällt. Aus (4) folgt für beliebige λ

$$(4a) \quad F = R(\lambda^2 R + 2\lambda Q + S) - (\lambda R + Q)^2.$$

Die Gleichung (4a) bedeutet, daß die Kurve F die Einhüllende des Systems

$$(5) \quad \lambda^2 R + 2\lambda Q + S = 0$$

ist. Jeder Kegelschnitt des Systems berührt die Kurve F in vier Punkten, die mit den vier Berührungspunkten der zwei Doppeltangenten t_1 und t'_1 auf einem Kegelschnitt $\lambda R + Q = 0$ liegen.

Aus (4) folgt:

$$(\lambda_1 - \lambda_2)^2 F = (\lambda_1^2 R + 2\lambda_1 Q + S)(\lambda_2^2 R + 2\lambda_2 Q + S) - [\lambda_1 \lambda_2 R + (\lambda_1 + \lambda_2)Q + S]^2;$$

daraus ergibt sich, daß je zwei Kegelschnitte des Systems die Kurve F in acht Punkten berühren, die auf einem Kegelschnitt liegen; insbesondere liegen die acht Berührungspunkte von je zwei zerfallenden Kegelschnitten des Systems (5) auf einem Kegelschnitt.

Will man R und S symmetrisch behandeln, so muß man die Formel (5) homogen machen, also sie durch

$$(6) \quad \lambda_1^2 R + 2\lambda_1 \lambda_2 Q + \lambda_2^2 S = 0$$

ersetzen.

Wir fragen nun, ob unter den Kegelschnitten des Systems (6) außer R noch weitere zerfallende Kegelschnitte vorkommen. Ein Kegelschnitt (6) zerfällt, wenn seine Diskriminante

$$(7) \quad \begin{vmatrix} \lambda_1^2 r_{00} + 2\lambda_1 \lambda_2 q_{00} + \lambda_2^2 s_{00} & \dots & \lambda_1^2 r_{02} + 2\lambda_1 \lambda_2 q_{02} + \lambda_2^2 s_{02} \\ \dots & \dots & \dots \\ \dots & \dots & \lambda_1^2 r_{22} + 2\lambda_1 \lambda_2 q_{22} + \lambda_2^2 s_{22} \end{vmatrix}$$

den Wert Null annimmt. Das ergibt eine Gleichung 6. Grades zur Bestimmung des Verhältnisses $\lambda_1 : \lambda_2$, von der wir zeigen wollen, daß sie keine mehrfachen Wurzeln haben kann.

Gesetzt etwa, $\lambda_1 = 0$ wäre eine mehrfache Wurzel, die Diskriminante (7) wäre also durch λ_1^2 teilbar. (Durch eine lineare Parametertransformation kann man jede andere Wurzel auf $\lambda_1 = 0$ zurückführen.) Wir wählen die Koordinaten so, daß der zur Wurzel $\lambda_1 = 0$ gehörige zerfallende Kegelschnitt die Gleichung $y_1 y_2 = 0$ hat [der Kegelschnitt S kann nicht in zwei zusammenfallende Geraden zerfallen, denn dann wäre wegen (4) jeder Schnittpunkt dieser Geraden mit dem Kegelschnitt Q ein Doppelpunkt der Kurve F , die doch doppeltpunktfrei sein sollte]; dann sind alle s_{k1} , mit Ausnahme von $s_{12} = s_{21}$, gleich Null. In der Determinante (7) sind jetzt die erste Zeile und die erste Spalte durch λ_1 teilbar; also alle Determinantenglieder durch λ_1^2 teilbar, außer dem einen

$$- 2 \lambda_1 \lambda_2 q_{00} \cdot \lambda_2^2 s_{12} \cdot \lambda_2^2 s_{21}.$$

Dieses ist nur dann durch λ_1^2 teilbar, wenn $q_{00} = 0$ ist, das heißt, wenn der Kegelschnitt Q durch den Punkt $(1, 0, 0)$, den Doppelpunkt des Kegelschnittes S , geht. Nach (4) wäre dann dieser Punkt auch Doppelpunkt von F , was unmöglich ist.

Also hat die obige Gleichung 6. Grades sechs verschiedene Wurzeln⁷⁾. Es folgt:

Das quadratische Kegelschnittsystem (5) enthält sechs zerfallende Kegelschnitte; jeder von ihnen zerfällt in zwei Doppeltangenten der Kurve F . Man nennt ein solches System von je sechs Doppeltangentenpaaren eine Steinersche Gruppe.

Ist $t_2 t'_2$ ein zweites Paar der durch $t_1 t'_1$ bestimmten Steinerschen Gruppe, so ist $t_2 t'_2$ im System $\lambda^2 R + 2 \lambda Q + S = 0$ enthalten, also gibt es ein solches λ , daß $t_2 t'_2 = \lambda^2 R + 2 \lambda Q + S$ wird. Man erhält dann, wenn statt $\lambda R + Q$ wieder Q geschrieben wird, nach (4a)

$$(8) \quad F = t_1 t'_1 t_2 t'_2 - Q^2.$$

Die obige Identität zeigt, daß das Doppeltangentenpaar $t_2 t'_2$ dieselbe Steinersche Gruppe bestimmt wie das Paar $t_1 t'_1$; also wird eine Steinersche Gruppe durch ein beliebiges der sechs Doppeltangentenpaare bestimmt. Da es $\frac{28 \cdot 27}{2}$ Paare von Doppeltangenten gibt, von denen je sechs eine Steinersche Gruppe bilden, so gibt es genau 63 Steinersche Gruppen.

Ist $t_3 t'_3$ ein drittes Paar derjenigen Steinerschen Gruppe, der auch $t_1 t'_1$ und $t_2 t'_2$ angehören, so ist $t_3 t'_3$ im System (5) enthalten, also $t_3 t'_3 = \lambda^2 R +$

⁷⁾ Dieser Beweis wurde einem unpublizierten Manuskript von Prof. van der Waerden entnommen.

+ 2 λQ + S . Daraus folgt: $Q = \frac{1}{2\lambda} (t_3 t'_3 - \lambda^2 t_1 t'_1 - t_2 t'_2)$. Nun kann man statt (8) auch schreiben:

$$\begin{aligned} F &= t_1 t'_1 (\lambda^2 t_1 t'_1 + 2 \lambda Q + t_2 t'_2) - (\lambda t_1 t'_1 + Q)^2 \\ &= t_1 t'_1 t_3 t'_3 - \left[\lambda t_1 t'_1 + \frac{1}{2\lambda} (t_3 t'_3 - \lambda^2 t_1 t'_1 - t_2 t'_2) \right]^2. \end{aligned}$$

Schreibt man nun für $\frac{t_3}{\lambda}$, λt_1 , $\frac{t_2}{\lambda}$ wieder t_3 , t_1 , t_2 , so erhält man

$$-4F = (t_3 t'_3 - t_2 t'_2 - t_1 t'_1)^2 - 4 t_1 t'_1 t_2 t'_2 = \begin{vmatrix} 0 & t_1 & t_2 & t_3 \\ t_1 & 0 & t'_3 & t'_2 \\ t_2 & t'_3 & 0 & t'_1 \\ t_3 & t'_2 & t'_1 & 0 \end{vmatrix}.$$

Es folgt, daß jedes geordnete System von sechs Doppeltangenten $t_1 t'_1$, $t_2 t'_2$, $t_3 t'_3$, die drei Paare einer Steinerschen Gruppe bilden, in eindeutiger Weise zu einer Hesseschen Erzeugung der Kurve vierter Ordnung führt, mit bestimmter Numerierung der vier ersten Basispunkte 1, 2, 3, 4, und zwar so, daß t_1 , t'_1 , t_2 , t'_2 , t_3 , t'_3 die Nummern 12, 34, 13, 24, 14, 23 erhalten; umgekehrt führt jede Hessesche Erzeugung mit vier numerierten Basispunkten 1, 2, 3, 4 zu einem solchen System $t_1 t'_1$, $t_2 t'_2$, $t_3 t'_3$. Da je drei Paare $3! \cdot 2^3 = 48$ solcher geordneten Systeme ergeben, so gibt es zu jeder Steinerschen Gruppe $48 \cdot 20 = 960$ Hessesche Erzeugungen. Die Anzahl der Steinerschen Gruppen ist 63, die Anzahl der geordneten Systeme $t_1 t'_1$, $t_2 t'_2$, $t_3 t'_3$ also $63 \cdot 960$, die Anzahl der Hesseschen Erzeugungen mit vier numerierten Basispunkten somit auch $63 \cdot 960$. Beachtet man die Willkür in der Bezeichnung der Basispunkte 1, 2, 3, 4 eines Quadrikenetzes, so folgt, daß es $\frac{63 \cdot 960}{8 \cdot 7 \cdot 6 \cdot 5} = 36$ Hessesche Erzeugungen einer vorgegebenen biquadratischen Kurve gibt.

1. 3. Sind $t_i t'_i$ und $t_j t'_j$ irgend zwei Doppeltangentenpaare einer Steinerschen Gruppe, so erhält man $F = t_i t'_i t_j t'_j - Q^2$, wobei Q der durch die acht Berührungspunkte gehende Kegelschnitt ist. In dieser Gleichung können die Rollen von $t_i t'_i$, $t_j t'_j$ beliebig vertauscht werden. Wenn also die Paare $t_i t'_i$, $t_j t'_j$ einer Steinerschen Gruppe angehören; so gehören auch $t_i t'_j$ und $t'_i t'_j$ sowie $t'_i t_j$ und $t_i t_j$ je einer Steinerschen Gruppe an. Man nennt ein solches Quadrupel, dessen acht Berührungspunkte auf einem Kegelschnitt liegen, ein syzygetisches Quadrupel. Weiter: je drei Doppeltangenten $t_i t'_i$, $t_j t'_j$, deren sechs Berührungspunkte auf einem Kegelschnitt liegen, lassen sich immer zu einem syzygetischen Quadrupel ergänzen. Denn setzt man wieder $R = t_i t'_i$ und nimmt für Q den Kegelschnitt durch die sechs Berührungspunkte, so gilt wie oben die Gleichung $F = RS - Q^2$, aus der folgt, daß der Kegelschnitt S die Kurve F in vier Punkten berührt, zu denen die zwei Berührungspunkte der Doppeltangente t_j gehören. Also enthält der Kegelschnitt die Gerade t_j als

Bestandteil. Setzt man nun $S = t_i t_j$, so geht diese Gleichung in $F = t_i t_j - Q^2$ über; es bilden also $t_i t_j$ ein syzygetisches Quadrupel. Ein solches Tripel von Doppeltangenten, deren Berührungspunkte auf einem Kegelschnitt liegen, nennt man ein syzygetisches Tripel; jedes andere Tripel heißt asyzygetisch.

1. 4. Nach 1. 1. hat jede Doppeltangente $\bar{j}k$ zwei „Pole“ E_j und E_k . Wir haben gesehen, daß jedes Doppeltangentenpaar eine Steinersche Gruppe bestimmt. Zwei Doppeltangenten haben aber entweder einen gemeinsamen Pol, wie $\bar{j}l$ und $\bar{k}l$, oder keinen, wie $\bar{j}k$ und $\bar{l}m$. Wir fragen nun: Welche Bezeichnung tragen die anderen fünf Doppeltangentenpaare, die mit $\bar{j}l \bar{k}l$ oder $\bar{j}k \bar{l}m$ eine Steinersche Gruppe bilden?

Wählt man irgend vier Basispunkte j, k, l, m des quadratischen Flächen-netzes als Ecken des Koordinatentetraeders, so sieht man nach (3), daß die Doppeltangenten $\bar{j}k \bar{l}m \bar{j}l \bar{m}k$ stets ein syzygetisches Quadrupel bilden. Somit gehören die Paare $\bar{j}k \bar{m}k$ und $\bar{j}l \bar{m}l$ oder die Paare $\bar{j}k \bar{l}m$ und $\bar{j}l \bar{m}k$ derselben Steinerschen Gruppe an. Daraus folgt, daß $\bar{1}k \bar{2}k$ ($k = 3, 4, 5, 6, 7, 8$) eine Steinersche Gruppe bilden. Es folgt weiter, daß die Tripel $\bar{1}3 \bar{2}3 \bar{1}2$, $\bar{1}2 \bar{2}3 \bar{4}5$ und $\bar{1}3 \bar{2}3 \bar{3}4$ asyzygetisch sind, denn in der durch $\bar{1}3 \bar{2}3$ bestimmten Steinerschen Gruppe kommen die Doppeltangenten $\bar{1}2, \bar{3}4$ und $\bar{4}5$ nicht vor. Allgemein ist das Tripel $\bar{1}k \bar{2}k \bar{1}2$, $\bar{1}l \bar{2}l \bar{k}l$ oder $\bar{1}l \bar{2}l \bar{j}k$ ($j, k, l = 3, 4, 5, 6, 7, 8$ und $j \neq k \neq l$) asyzygetisch.

Bezeichnen wir irgendeine Doppeltangente mit einem Strich: |, zwei Doppeltangenten, die einen gemeinschaftlichen Pol haben, mit \wedge , sonst mit \parallel und entsprechend für Tripel und Quadrupel, so können wir das eben Gesagte auch so ausdrücken: Die Quadrupel \square und daher die Tripel \sqcap sind syzygetisch, die Tripel \triangle , \vee und \vee dagegen asyzygetisch.

Das Paar $\bar{1}2 \bar{3}4$ definiert eine Steinersche Gruppe, welche nach dem obigen keine Doppeltangenten $\bar{k}l$ ($k = 1, 2, 3, 4; l = 5, 6, 7, 8$) enthalten kann. Also kann die Gruppe nur aus den Paaren $\bar{1}2 \bar{3}4, \bar{1}3 \bar{4}2, \bar{1}4 \bar{2}3, \bar{5}6 \bar{7}8, \bar{5}7 \bar{6}8, \bar{5}8 \bar{6}7$ bestehen. Demnach sind die Quadrupel $\parallel\parallel\parallel\parallel$ und die Tripel $\parallel\parallel$ syzygetisch.

Hieraus erkennt man, daß es zwei Arten von Steinerschen Gruppen gibt:

Bei der ersten Art haben die Doppeltangenten je eines Paares der Gruppe einen Pol gemeinsam und die sechs Doppeltangentenpaare zwei Pole E_j und E_k gemeinsam; man bezeichnet eine solche Steinersche Gruppe mit $\{jk\}$. Ihre Anzahl beträgt 28. Beispiel: $\bar{1}3 \bar{2}3, \bar{1}4 \bar{2}4, \bar{1}5 \bar{2}5, \bar{1}6 \bar{2}6, \bar{1}7 \bar{2}7, \bar{1}8 \bar{2}8$.

Bei der zweiten Art haben je zwei Doppeltangenten eines Paares irgendeiner Gruppe keinen Pol gemeinsam und je zwei Doppeltangentenpaare entweder vier Pole gemeinsam oder gar keinen. Wenn sie vier Pole E_j, E_k, E_l, E_m gemeinsam haben, dann gehört noch das dritte Doppeltangentenpaar

dazu, welches auch die vier Pole hat; die drei übrigen Doppeltangentenpaare haben die vier anderen Pole E_p, E_q, E_r, E_s gemeinsam. Man bezeichnet eine solche Steinersche Gruppe mit $\{jklm\}$ oder $\{pqrs\}$; ihre Anzahl beträgt 35. Beispiel: 12 34, 13 24, 14 23, 56 73, 57 68, 58 67.

Wählt man aus irgendeiner Steinerschen Gruppe drei beliebige Doppeltangentenpaare, hernach aus jedem Paar eine beliebige Doppeltangente, so bilden diese ein asyzygetisches Tripel, da ein solches Tripel nur das Symbol Δ , ψ oder ψ haben kann.

1. 5. Seien $t_1 t_2 t_3 t_4$ und $u_1 u_2 u_3 u_4$ irgend zwei syzygetische Quadrupel, dann erhält man

$$\begin{aligned} F &= t_1 t_2 t_3 t_4 - Q^2 \\ &= u_1 u_2 u_3 u_4 - Q'^2, \end{aligned}$$

wobei Q und Q' zwei Kegelschnitte sind. Daraus folgt

$$t_1 t_2 t_3 t_4 - u_1 u_2 u_3 u_4 = (Q + Q')(Q - Q').$$

Die obige Identität zeigt, daß die 16 Schnittpunkte von t und u auf zwei Kegelschnitten $Q + Q'$ und $Q - Q'$ liegen. Wenn die zwei Quadrupel eine Doppeltangente gemeinsam haben, z. B. $t_4 = u_4$, dann ist $t_4 (t_1 t_2 t_3 - u_1 u_2 u_3) = (Q + Q')(Q - Q')$. So zerfällt entweder $Q + Q'$ oder $Q - Q'$ in zwei Geraden t_4 und t'_4 . In jedem der beiden Fälle ergibt sich, daß die 9 Schnittpunkte von $t_1 t_2 t_3$ und $u_1 u_2 u_3$ auf einem Kegelschnitt und einer Geraden t'_4 liegen; das heißt, daß die sechs Doppeltangenten ein Pascalsches Sechseck bilden. Es folgt:

Die 16 Schnittpunkte von je zwei syzygetischen Quadrupeln liegen auf zwei Kegelschnitten. Wenn die beiden eine Doppeltangente gemeinsam haben, so bilden die übrigen sechs ein Pascalsches Sechseck.

1. 6. Wir schreiben das Kegelschnittssystem $2^2 R + 2 \lambda Q + S = 0$ in Linienkoordinaten $\sum A_{ik} u_i u_k = 0$, wobei $A_{ik} = \alpha_{ik} + \beta_{ik} \lambda + \gamma_{ik} \lambda^2 + \delta_{ik} \lambda^3 + \varepsilon_{ik} \lambda^4$ ist. Sei b_{ik} eine Lösung der fünf Gleichungen

$$\sum \alpha_{ik} b_{ik} = 0,$$

$$\sum \beta_{ik} b_{ik} = 0,$$

$$\sum \gamma_{ik} b_{ik} = 0,$$

$$\sum \delta_{ik} b_{ik} = 0,$$

$$\sum \varepsilon_{ik} b_{ik} = 0,$$

dann ist

$$(9) \quad \sum A_{ik} b_{ik} = 0$$

identisch in λ . Weiter sei x_i der Doppelpunkt eines zerfallenden Kegelschnittes des Systems, dann ist $\sum A_{ik} u_i u_k = (\sum x_i u_i)^2$. Daraus folgt:

$$(10) \quad A_{ik} = x_i x_k.$$

Durch Einsetzen von (10) in (9) erhalten wir

$$\sum x_i x_k b_{ik} = 0,$$

das heißt, der Doppelpunkt x_i liegt auf dem Kegelschnitt

$$\sum b_{ik} x_i x_k = 0.$$

Es folgt:

Die sechs Schnittpunkte x_i der sechs Doppeltangentenpaare einer Steiner'schen Gruppe liegen auf einem Kegelschnitt⁸⁾.

II. Punktgruppen auf einer Kurve vierter Ordnung.

2. 1. Betrachtet man je endlich viele Punkte $P_1 P_2 P_3 \dots P_n$ auf einer Kurve zusammen als eine Punktgruppe und bezeichnet diese mit D ; betrachtet man weiter zwei Punktgruppen $D' D''$ zusammen genommen als ihr Produkt und bezeichnet es mit $D = D' D''$, dann sagt der Restsatz aus:

Wenn $D' D''$ von einer Kurve $(m + n)$ -ter Ordnung, D' von einer Kurve m -ter Ordnung ausgeschnitten wird, so wird D'' von einer Kurve n -ter Ordnung ausgeschnitten⁹⁾.

Wenn es eine Kurve K von m -ter Ordnung gibt, die die Punktgruppe DC ausschneidet, und eine Kurve K' von m -ter Ordnung, die die Punktgruppe $D'C$ ausschneidet, so heißen D und D' äquivalent, und man schreibt $D \sim D'$. Dann folgt leicht:

- (a) $D \sim D$;
- (b) wenn $D \sim D'$, so $D' \sim D$;
- (c) wenn $D \sim D'$, $D' \sim D''$, so $D \sim D''$;
- (d) wenn $D \sim D'$, $E \sim E'$, so $DE \sim D'E'$.

So sind alle Punktgruppen, die zu einer festen äquivalent sind, auch untereinander äquivalent und bilden eine Klasse von Punktgruppen, die man mit K_D bezeichnet.

Aus zwei Klassen K_D und K_E bildet man alle Punktgruppen $D_i E_j$, die untereinander wieder äquivalent sind. Man findet also eine neue Klasse und bezeichnet sie mit $K_D K_E = K_{DE}$, wonach $D_i E_j \in K_{DE}$ ist.

⁸⁾ Siehe E. Caporali, Memorie di geometria, Napoli 1888, S. 364.

⁹⁾ Vgl. den „Satz vom Doppelpunktsdivisor“ bei v. d. Waerden, Einführung in die algebraische Geometrie, Berlin 1939, S. 215. In unserem Fall ist der Doppelpunktsdivisor D leer, da die Grundkurve keine Doppelpunkte hat.

Sind π und π' irgend zwei Doppeltangentenpaare einer Steinerschen Gruppe, ist weiter α die Gruppe der vier Berührungspunkte von π und β die Gruppe der vier Berührungspunkte von π' , dann liegen $\alpha\beta$ auf einem Kegelschnitt Q und $\alpha\alpha$ auf dem Kegelschnitt π , also sind α und β äquivalent. Umgekehrt sei $\gamma \sim \delta$, wobei γ die Berührungspunktgruppe des Doppeltangentenpaares π_1 und δ die Berührungspunktgruppe des Doppeltangentenpaares π_2 ist, so ist auch $\gamma\gamma \sim \gamma\delta$, und es gibt eine Kurve K m -ter Ordnung, die die Grundkurve in $\gamma\gamma R$ schneidet und eine Kurve K' m -ter Ordnung, die die Grundkurve in $\gamma\delta R$ schneidet. Da $\gamma\gamma$ auf einem Kegelschnitt liegen, wird nach dem Restsatz R von einer Kurve $(m-2)$ -ter Ordnung ausgeschnitten. Da weiter $\gamma\delta R$ von der Kurve K' ausgeschnitten wird, so wird, wieder nach dem Restsatz, $\gamma\delta$ von einem Kegelschnitt ausgeschnitten. Somit gehören π_1 und π_2 derselben Steinerschen Gruppe an. Es folgt:

Je sechs Punktgruppen auf einer Kurve vierter Ordnung, von denen jede aus den vier Berührungspunkten eines Doppeltangentenpaares einer Steinerschen Gruppe besteht, gehören einer Klasse an und umgekehrt.

Man bezeichnet die Klasse mit $[jk]$ oder $[jklm]$, je nachdem das Doppeltangentenpaar, dessen Berührungspunkte die Punktgruppe bilden, der Steinerschen Gruppe $\{jk\}$ oder $\{jklm\}$ angehört. Man bezeichnet die Gruppe der Berührungspunkte des Doppeltangentenpaares $\bar{j}\bar{k}\bar{l}\bar{m}$ mit $(jklm)$, die Gruppe der Berührungspunkte der Doppeltangente $\bar{j}\bar{k}$ mit (jk) . Dann ergibt sich:

1. $(jk) = (kj)$;
2. $(jklk) \sim (jmk m)$;
3. $(jklm) \sim (jlmk) \sim (jmk l) \sim (pqrs) \sim (prsq) \sim (psqr)$;
4. $[jklm] \equiv [jlmk] \equiv [jmk l] \equiv [pqrs] \equiv [prsq] \equiv [psqr]$.

2. 2. Lemma. Indem man aus zwei Steinerschen Gruppen je ein Doppeltangentenpaar nimmt, kann man immer zwei solche Doppeltangentenpaare erhalten, die eine zweimal vorkommende Doppeltangente enthalten. Und die zwei übrigen Doppeltangenten bilden ein Doppeltangentenpaar, welches zu einer dritten Steinerschen Gruppe gehört, und zwar erhält man das Symbol $\{jk\}$ oder $\{jklm\}$ dieser dritten Gruppe, indem man die Symbole der ersten beiden Gruppen nebeneinander schreibt und die zweimal vorkommenden Ziffern wegläßt, oder auch, indem man die einmal vorkommenden Ziffern wegläßt und die nicht oder zweimal vorkommenden aufschreibt. Man nennt sie die Produktgruppen der beiden Gruppen.

Beweis. Wenn die zwei Steinerschen Gruppen beide der ersten Art der Steinerschen Gruppen angehören, so gibt es zwei Fälle: $\{jk\}$, $\{lm\}$ und $\{jk\}$, $\{jl\}$. Von $\{jk\}$, $\{lm\}$ enthält $\{jk\}$ das Paar $\bar{j}\bar{l}\bar{k}\bar{l}$ und $\{lm\}$ das Paar $\bar{l}\bar{j}\bar{m}\bar{j}$; also enthalten die zwei Paare eine zweimal vorkommende Doppeltangente $\bar{j}\bar{l}$

und das der Steinerschen Gruppe $\{lmjk\}$ angehörende Doppeltangentenpaar $\overline{klm}j$. Von $\{jk\}$, $\{jl\}$ enthält $\{jk\}$ das Paar $\overline{jl}kl$ und $\{jl\}$ das Paar $\overline{jk}lk$; also enthalten die zwei Paare eine zweimal vorkommende Doppeltangente \overline{kl} und das der Steinerschen Gruppe $\{kl\}$ angehörende Doppeltangentenpaar $\overline{kl}j$. Wenn die zwei Steinerschen Gruppen beide zu der zweiten Art der Steinerschen Gruppen gehören, so gibt es zwei Fälle: $\{jklm\}$, $\{jkpq\}$ und $\{jklm\}$, $\{jklp\}$. Von $\{jklm\}$, $\{jkpq\}$ enthält $\{jklm\}$ das Paar $\overline{jk}lm$ und $\{jkpq\}$ das Paar $\overline{jk}pq$; also enthalten die zwei Paare eine zweimal vorkommende Doppeltangente \overline{jk} und das der Steinerschen Gruppe $\{lm pq\}$ angehörende Doppeltangentenpaar $\overline{lm}pq$. Von $\{jklm\}$, $\{jklp\}$ enthält $\{jklm\}$ das Paar $\overline{jk}lm$ und $\{jklp\}$ das Paar $\overline{jk}lp$; also enthalten die zwei Paare eine zweimal vorkommende Doppeltangente \overline{jk} und das zu der Steinerschen Gruppe $\{mp\}$ gehörige Doppeltangentenpaar $\overline{lm}lp$.

Wenn eine der zwei Steinerschen Gruppen der ersten Art und die andere der zweiten Art angehört, dann gibt es zwei Fälle: $\{jklm\}$, $\{jk\}$ und $\{jklm\}$, $\{jp\}$. Von $\{jklm\}$, $\{jk\}$ enthält $\{jklm\}$ das Paar $\overline{jl}km$ und $\{jk\}$ das Paar $\overline{jl}kl$; also enthalten die zwei Paare eine zweimal vorkommende Doppeltangente \overline{jl} und das zu der Steinerschen Gruppe $\{lm\}$ gehörende Doppeltangentenpaar $\overline{kl}km$. Von $\{jklm\}$, $\{jp\}$ enthält $\{jklm\}$ das Paar $\overline{jk}lm$ und $\{jp\}$ das Paar $\overline{jk}pk$; also enthalten die zwei Paare eine zweimal vorkommende Doppeltangente \overline{jk} und das der Steinerschen Gruppe $\{klmp\}$ angehörende Doppeltangentenpaar $\overline{lm}pk$.

Die so erklärte Multiplikation ist offenbar kommutativ und assoziativ, denn die Produktgruppe von drei oder mehr Gruppen wird immer so erhalten, daß die Symbole dieser Gruppen nebeneinander geschrieben werden und dann entweder die Ziffern, die eine gerade Anzahl Male vorkommen, weggelassen und die anderen beibehalten werden, oder umgekehrt.

2. 3. Nun untersuchen wir die Punktgruppen auf einer Kurve vierter Ordnung.

Kriterium. Die Punktgruppe $P = P_1 P_2 \dots P_{4n}$ wird dann und nur dann von einer Kurve n -ter Ordnung ausgeschnitten, wenn $P \sim (12)^{2n}$ ist.

Das Kriterium folgt unmittelbar aus dem Restsatz, weil die $4n$ Punkte $(12)^{2n}$ immer von einer Kurve n -ter Ordnung, nämlich der n mal gezählten Geraden $\overline{12}$, ausgeschnitten werden.

Hauptsatz. (1) Je $n-1$ ($2 < n < 8$) Steinersche Gruppen, von denen keine die Produktgruppe von m ($m < n-1$) aus den übrigen ist, bilden mit ihrer Produktgruppe ein System $S[n]$, und zwar gibt es

$$\frac{63(63-1)\left(63-2-\binom{2}{2}\right) \dots \left(63-(n-2)-\binom{n-2}{2}-\binom{n-2}{3}\right) \dots \binom{n-2}{n-2}}{n!}$$

Systeme $S[n]$.

(2) Wählt man aus jeder der n Steinerschen Gruppen ein beliebiges Doppeltangentenpaar, so werden die $4n$ Berührungspunkte von einer Kurve n -ter Ordnung ausgeschnitten.

(3) Werden $4n$ Berührungspunkte von n Doppeltangentenpaaren von einer Kurve n -ter Ordnung ausgeschnitten, so gehören die n Doppeltangentenpaare je einer Steinerschen Gruppe eines Systems $S[n]$ an.

Beweis. (1) Für die erste Gruppe kann man jede der 63 Steinerschen Gruppen, für die zweite jede der übrigen $63 - 1$ wählen. Für die dritte aber kann man nur jede von $63 - 2 - \binom{2}{2}$ Steinerschen Gruppen wählen; weil nämlich je zwei der gewählten Steinerschen Gruppen eine Produktgruppe haben, muß man diese auch abrechnen. Allgemeiner kann man für die r -te Gruppe nur jede von den übrigen $63 - (r-1) - \binom{r-1}{2} - \binom{r-1}{3} \dots - \binom{r-1}{r-1}$ Steinerschen Gruppen wählen; deshalb gibt es im Ganzen

$$\frac{63(63-1)\left(63-2-\binom{2}{2}\right) \dots \left(63-(n-2)-\binom{n-2}{2} \dots -\binom{n-2}{n-2}\right)}{n!}$$

Systeme $S[n]$.

(2) Sei π_i ein beliebiges Doppeltangentenpaar der i -ten Gruppe und π_{P_m} ein beliebiges Doppeltangentenpaar der Produktgruppe $\{P_m\}$, dann ist

$$(\pi_1)(\pi_2) \sim (jk)^2(\pi_{P_2}),$$

weil $(12)^2 \sim (13)^2 \dots \sim (18)^2 \sim (23)^2 \dots \sim (28)^2 \dots \sim (78)^2$ ist, folgt daraus

$$(\pi_1)(\pi_2) \sim (12)^2(\pi_{P_2}),$$

$$(\pi_1)(\pi_2)(\pi_3) \sim (12)^2(\pi_{P_2})(\pi_3) \sim (12)^2(12)^2(\pi_{P_3}) \sim (12)^4(\pi_{P_3});$$

und

$$(\pi_1)(\pi_2) \dots (\pi_{n-1}) \sim (12)^{2(n-2)}(\pi_{P_{n-1}});$$

also

$$(\pi_1)(\pi_2) \dots (\pi_{n-1})(\pi_{P_{n-1}}) \sim (12)^{2(n-2)}(\pi_{P_{n-1}})(\pi_{P_{n-1}}) \sim (12)^{2n};$$

deshalb werden die $4n$ Berührungspunkte von einer Kurve n -ter Ordnung ausgeschnitten.

(3) Da jedes Doppeltangentenpaar einer Steinerschen Gruppe angehört, so gehört jedes Paar π_i der n Doppeltangentenpaare $\pi_1 \pi_2 \pi_3 \dots \pi_n$ zu einer bestimmten Steinerschen Gruppe g_i . Sei $g' = g_1 g_2 \dots g_{n-1}$ und π' irgendein Doppeltangentenpaar der Steinerschen Gruppe g' , so werden $(\pi_1 \pi_2 \pi_3 \dots \pi_{n-1} \pi')$ nach (2) von einer Kurve n -ter Ordnung K ausgeschnitten. Nach der Voraussetzung werden $(\pi_1 \pi_2 \pi_3 \dots \pi_{n-1} \pi_n)$ von einer Kurve n -ter Ordnung K' ausgeschnitten, also ist $(\pi_n) \sim (\pi')$ und $(\pi_n \pi')$ werden von einem Kegelschnitt ausgeschnitten. Mithin gilt $g_n \equiv g'$, somit bilden g_1, g_2, \dots, g_n ein System $S[n]$.

Def. Ein 2 n -tupel von Doppeltangenten, welches aus n Doppeltangentenpaaren aus den n Steinerschen Gruppen eines Systems $S[n]$ besteht, heißt syzygetisch.

2. 4. Wendet man den Hauptsatz auf die Systeme $S[3]$ an, so findet man vier Arten:

1. Irgendein System der ersten Art wird von 3 Basispunkten definiert, etwa $\{jk\}$, $\{kl\}$, $\{lj\}$. Die Anzahl der Systeme von dieser Art ist 56;

z. B. $\{12\}$: $\overline{13} \overline{23}$, $\overline{14} \overline{24}$, $\overline{15} \overline{25}$, $\overline{16} \overline{26}$, $\overline{17} \overline{27}$, $\overline{18} \overline{28}$;
 $\{23\}$: $\overline{21} \overline{31}$, $\overline{24} \overline{34}$, $\overline{25} \overline{35}$, $\overline{26} \overline{36}$, $\overline{27} \overline{37}$, $\overline{28} \overline{38}$;
 $\{31\}$: $\overline{32} \overline{12}$, $\overline{34} \overline{14}$, $\overline{35} \overline{15}$, $\overline{36} \overline{16}$, $\overline{37} \overline{17}$, $\overline{38} \overline{18}$.

2. Irgendein System der zweiten Art wird von 4 Basispunkten definiert, etwa $\{jklm\}$, $\{jk\}$, $\{lm\}$. Die Anzahl der Systeme von dieser Art ist 105;

z. B. $\{1234\}$: $\overline{12} \overline{34}$, $\overline{13} \overline{24}$, $\overline{14} \overline{23}$, $\overline{56} \overline{78}$, $\overline{57} \overline{68}$, $\overline{58} \overline{67}$;
 $\{12\}$: $\overline{13} \overline{23}$, $\overline{14} \overline{24}$, $\overline{15} \overline{25}$, $\overline{16} \overline{26}$, $\overline{17} \overline{27}$, $\overline{18} \overline{28}$;
 $\{34\}$: $\overline{31} \overline{41}$, $\overline{32} \overline{42}$, $\overline{35} \overline{45}$, $\overline{36} \overline{46}$, $\overline{37} \overline{47}$, $\overline{38} \overline{48}$.

3. Irgendein System der dritten Art wird von 5 Basispunkten definiert, etwa $\{jklm\}$, $\{jklp\}$, $\{mp\}$. Die Anzahl der Systeme von dieser Art ist 280;

z. B. $\{1234\}$: $\overline{12} \overline{34}$, $\overline{13} \overline{24}$, $\overline{14} \overline{23}$, $\overline{56} \overline{78}$, $\overline{57} \overline{68}$, $\overline{58} \overline{67}$;
 $\{1235\}$: $\overline{12} \overline{35}$, $\overline{13} \overline{25}$, $\overline{15} \overline{23}$, $\overline{46} \overline{78}$, $\overline{47} \overline{68}$, $\overline{48} \overline{67}$;
 $\{45\}$: $\overline{41} \overline{51}$, $\overline{42} \overline{52}$, $\overline{43} \overline{53}$, $\overline{46} \overline{56}$, $\overline{47} \overline{57}$, $\overline{48} \overline{58}$.

4. Irgendein System der vierten Art wird von 6 Basispunkten definiert, etwa $\{lmpq\}$, $\{pqjk\}$, $\{jklm\}$. Die Anzahl der Systeme von dieser Art ist 210;

z. B. $\{1234\}$: $\overline{12} \overline{34}$, $\overline{13} \overline{24}$, $\overline{14} \overline{23}$, $\overline{56} \overline{78}$, $\overline{57} \overline{68}$, $\overline{58} \overline{67}$;
 $\{3456\}$: $\overline{34} \overline{56}$, $\overline{35} \overline{46}$, $\overline{36} \overline{54}$, $\overline{12} \overline{78}$, $\overline{17} \overline{28}$, $\overline{18} \overline{27}$;
 $\{5612\}$: $\overline{56} \overline{12}$, $\overline{51} \overline{62}$, $\overline{52} \overline{61}$, $\overline{34} \overline{78}$, $\overline{37} \overline{48}$, $\overline{38} \overline{47}$.

Jedes System der zweiten oder vierten Art umfaßt alle die 28 Doppeltangenten. Man bezeichnet diese Systeme zusammen mit $S_1[3]$. Ein System der ersten oder dritten Art aber umfaßt nur 18 verschiedene Doppeltangenten. Man bezeichnet diese Systeme mit $S_2[3]$. Daraus erkennt man, daß die drei Steinerschen Gruppen eines jeden Systems $S_1[3]$ vier solche Doppeltangenten, die in jeder Steinerschen Gruppe zwei Doppeltangentenpaare bilden, gemeinsam haben; die drei Steinerschen Gruppen eines Systems $S_2[3]$ haben keine gemeinsamen Doppeltangenten.

Weiter, wenn drei Doppeltangentenpaare, die man aus den drei Gruppen des Systems $S[3]$ wählt, drei zweimal vorkommende Doppeltangenten umfassen, so zerfällt die Kurve Q_3 , die durch die 12 Berührungspunkte geht, in drei Geraden; wenn sie eine zweimal vorkommende Doppeltangente und

vier andere, die zwei Doppeltangentenpaare einer gewissen Steinerschen Gruppe bilden, umfassen, dann zerfällt Q_3 in eine Gerade und einen Kegelschnitt. Man bezeichnet jede der vier gemeinsamen Doppeltangenten eines Systems S_1 [3] mit t_0 und jeden durch die acht Berührungspunkte dieser vier Doppeltangenten gehenden Kegelschnitt mit Q_0 , die 24 übrigen Doppeltangenten mit t , den durch die 8 Berührungspunkte von irgend vier Doppeltangenten t oder t und t_0 gehenden Kegelschnitt mit Q . Durch alleinige Kombination findet man folgendes:

1. In jedem System S_1 [3] bestehen von den 216 Kurven Q_3
 - a) 4 aus je drei der vier t_0 ,
 - b) 4 aus $Q_0 + t_0$,
 - c) 4 aus $Q + t_0$,
 - d) die übrigen 160 sind eigentliche Q_3 .
2. In jedem System S_2 [3] bestehen von den 216 Kurven Q_3
 - e) 6 aus 3 t ,
 - f) 90 aus $Q + t$,
 - g) die übrigen 120 sind die eigentlichen Kurven Q_3 .

An verschiedenen Kurven Q_3 gibt es im ganzen

3276, die aus je 3 t .

1260, die aus $Q_0 + t_0$.

7560, die aus $Q + t$ bestehen, und

6048 eigentliche Kurven Q_3 .

Seien $t_1 t_2 \dots t_6$ die 6 Doppeltangenten eines syzygetischen 6-tupels und Q_3 die durch die 12 Berührungspunkte gehende Kurve dritter Ordnung, dann definiert $\lambda Q_3^2 + t_1 t_2 \dots t_6 = 0$ ein Büschel von Kurven sechster Ordnung. Jede Kurve des Büschels berührt die Kurve F in denselben 12 Punkten. Wählt man also λ so, daß die Kurve $\lambda Q_3^2 + t_1 t_2 \dots t_6 = 0$ noch durch einen weiteren Punkt von F geht, so hat diese Kurve mit F 25 Punkte gemeinsam; F ist dann als Bestandteil in ihr enthalten: man hat daher die Identität

$$\lambda Q_3^2 + t_1 t_2 \dots t_6 = F P.$$

oder wenn man λQ_3^2 als $-Q_3^2$ bezeichnet,

$$F P = t_1 t_2 \dots t_6 - Q_3^2.$$

wobei $P = 0$ ein Kegelschnitt ist.

Die obige Identität bedeutet, daß die zerfallende Kurve sechster Ordnung $F P = 0$ die Einhüllende des Systems $\lambda^2 t_1 t_2 t_3 + 2 \lambda Q_3 + t_4 t_5 t_6 = 0$ ist, und $t_1 t_2 t_3$ sowie $t_4 t_5 t_6$ sind zwei zerfallende Kurven des Systems. Weiter: Jede

Doppeltangente t_i berührt den Kegelschnitt P , und die Kurve Q_3 geht noch durch die 6 Berührungspunkte von P mit t_i . Nach einem bekannten Satze folgt:

Die 6 Ecken zweier Dreiseite $t_1 t_2 t_3$ und $t_4 t_5 t_6$, wobei $t_1 t_2 \dots t_6$ die Doppeltangenten eines syzygetischen 6-tupels sind, liegen auf einem Kegelschnitt, und zwar definiert ein syzygetisches 6-tupel 10 solche Kegelschnitte.

2. 5. Wendet man den Hauptsatz auf die Systeme $S[4]$ an, so findet man 11 Arten:

Irgendein System der ersten oder zweiten Art wird von 4 Basispunkten definiert, und zwar gilt folgendes:

1. Bei der ersten Art sind die Systeme so beschaffen: $\{jk\}$, $\{kl\}$, $\{lm\}$, $\{mj\}$. Die Anzahl der Systeme von dieser Art beträgt $\binom{8}{4} \cdot 3! \cdot \frac{1}{2} = 210$.

2. Bei der zweiten Art ist irgendein System so beschaffen: $\{jk\}$, $\{jl\}$, $\{jm\}$, $\{jklm\}$. Die Anzahl der Systeme von dieser Art beträgt $\binom{8}{4} \cdot 4 = 280$.

Irgendein System der dritten, vierten oder fünften Art wird von 5 Basispunkten definiert, und zwar:

3. Bei der dritten Art ist irgendein System so beschaffen: $\{jklm\}$, $\{jk\}$, $\{lp\}$, $\{mp\}$. Die Anzahl der Systeme von dieser Art beträgt $\binom{8}{5} \cdot \binom{5}{4} \cdot \binom{4}{2} = 1680$.

4. Bei der vierten Art ist irgendein System so beschaffen: $\{jklm\}$, $\{klmp\}$, $\{jl\}$, $\{lp\}$. Die Anzahl der Systeme von dieser Art beträgt $\binom{8}{5} \cdot 5 \cdot \binom{4}{2} = 1680$.

5. Bei der fünften Art ist irgendein System so beschaffen: $\{jklm\}$, $\{jklp\}$, $\{jkmp\}$, $\{jk\}$. Die Anzahl der Systeme von dieser Art beträgt $\binom{8}{5} \cdot \binom{5}{2} = 560$.

Irgendein System der sechsten, siebenten, achten oder neunten Art wird von 6 Basispunkten definiert, und zwar:

6. Bei der sechsten Art ist irgendein System so beschaffen: $\{jklm\}$, $\{jkpq\}$, $\{lp\}$, $\{mq\}$. Die Anzahl der Systeme von dieser Art beträgt $\binom{8}{6} \cdot \binom{6}{4} \cdot \binom{4}{2} \cdot \frac{1}{2} = 1260$.

7. Bei der siebenten Art ist irgendein System so beschaffen: $\{jk\}$, $\{lm\}$, $\{jkpq\}$, $\{lmpq\}$. Die Anzahl der Systeme von dieser Art beträgt $\binom{8}{6} \cdot \binom{6}{2} \cdot 3 \cdot \frac{1}{2} = 630$.

8. Bei der achten Art ist irgendein System so beschaffen: $\{jk\}$, $\{jklm\}$, $\{jlpq\}$, $\{jmpq\}$. Die Anzahl der Systeme von dieser Art beträgt $\binom{8}{6} \cdot 6 \cdot 5 \cdot \binom{4}{2} \cdot \frac{1}{2} = 2520$.

9. Bei der neunten Art ist irgendein System so beschaffen: $\{jklm\}, \{klmp\}, \{lmpq\}, \{jlmq\}$. Die Anzahl der Systeme von dieser Art beträgt $\binom{8}{6} \cdot \binom{6}{2} \cdot 3 \cdot \frac{1}{2} = 630$.

10. Irgendein System der zehnten Art wird von 7 Basispunkten definiert, etwa $\{jklm\}, \{jkpq\}, \{jlpq\}, \{jmqr\}$. Die Anzahl der Systeme von dieser Art beträgt $\frac{\binom{8}{7} \cdot 7 \cdot \binom{6}{3} \cdot 3!}{4 \cdot 8} = 210$.

11. Irgendein System der elften Art wird von 8 Basispunkten definiert, etwa $\{jk\}, \{lm\}, \{pq\}, \{rs\}$. Die Anzahl der Systeme von dieser Art beträgt $8! \cdot \frac{1}{4} \cdot \frac{1}{2^4} = 105$.

Im ganzen gibt es 9675 Systeme S [4].

Wählt man wie bei S [3] aus jeder Steinerschen Gruppe eines Systems S [4] ein beliebiges Doppeltangentenpaar, so geht durch die 16 Berührungspunkte eine Kurve Q_4 . Außer dieser gibt es aber noch unendlich viele, denn die Q_4 gehört einer linearen Schar $\lambda_0 F + \lambda_1 Q_4 = 0$ an, wobei F die Grundkurve ist. Und zwar ist diese Schar irreduzibel. Da nämlich erstens F und Q_4 keinen gemeinsamen Bestandteil haben, zweitens F irreduzibel ist, so ist nach einem bekannten Bertinischen Satz¹⁰⁾ diese Schar irreduzibel.

a) Wenn die vier Doppeltangentenpaare vier zweimal vorkommende Doppeltangenten umfassen, dann zerfällt eine Kurve Q_4 in $t_1 t_2 t_3 t_4$.

b) Wenn die vier Doppeltangentenpaare zwei zweimal vorkommende Doppeltangenten umfassen, dann zerfällt eine Kurve Q_4 in $t_1 t_2 Q$.

c) Wenn die vier Doppeltangentenpaare eine zweimal vorkommende Doppeltangente und sechs andere syzygetische Doppeltangenten umfassen, dann zerfällt eine Kurve Q_4 in t und Q_3 .

d) Wenn die vier Doppeltangentenpaare zwei syzygetische Doppeltangentenquadrupel umfassen, dann zerfällt eine Kurve Q_4 in Q und Q' .

Offenbar tritt in jedem Falle nur eine solche zerfallende Kurve auf. So gibt es in S [4] nur endlich viele in t, Q, Q_3 zerfallende Kurven Q_4 . Dabei sieht man auch, daß, wenn eine Kurve Q_4 nach Fall a) zerfällt, alle Kurven der Schar die Grundkurve in denselben acht Punkten berühren. Wenn eine Kurve Q_4 nach Fall b) zerfällt, so berühren alle Kurven der Schar die Grundkurve in denselben vier Punkten, usw.

Es seien $t_1 t_2 \dots t_8$ acht Doppeltangenten eines syzygetischen 8-tupels, weiter sei Q_4 eine durch die 16 Berührungspunkte gehende Kurve vierter Ordnung, dann definiert $\lambda Q_4^2 + t_1 t_2 \dots t_8 = 0$ ein Büschel von Kurven

¹⁰⁾ Siehe etwa B. L. v. d. Waerden, Einführung in die algebraische Geometrie, Berlin 1939, S. 204.

achter Ordnung. Jede Kurve im Büschel berührt die Kurve F in denselben 16 Punkten. Wählt man also λ so, daß die Kurve $\lambda Q_4^2 + t_1 t_2 \dots t_8 = 0$ noch durch einen Punkt von F geht, so hat diese Kurve mit F 33 Punkte gemeinsam; also ist F als Bestandteil in ihr enthalten. Man erhält somit die Identität $\lambda Q_4^2 + t_1 t_2 \dots t_8 = F P_4$, wobei P_4 eine Kurve vierter Ordnung ist. Diese Identität zeigt, daß jede Doppeltangente t_i die Kurve P_4 auch in zwei Punkten berührt und daß die Kurve Q_4 auch durch die 16 Berührungspunkte von P_4 mit t_i geht.

2. 6. Wählt man ebenso aus jeder Steinerschen Gruppe eines Systems $S[n]$ für $4 < n < 8$ ein beliebiges Doppeltangentenpaar, so gehen durch die $4n$ Berührungspunkte unendlich viele Kurven Q_n , die eine irreduzibel lineare Schar bilden. Jede Kurve Q_n definiert nämlich eine Linearschar

$$\lambda_0 Q_n + (\lambda_1 x_0^{n-4} + \lambda_2 x_0^{n-5} x_1 + \dots + \lambda_n x_2^{n-4}) F = 0.$$

Wenn Q_n die Grundkurve nicht als Bestandteil enthält, so haben Q_n und $x_0^{n-4} F$ und $x_1^{n-4} F$ keinen gemeinsamen Bestandteil. Wäre nun die Schar reduzibel, so müßten nach dem Satz von Bertini alle Kurven der Schar, insbesondere also alle Kurven $P_{n-4} F = 0$ in Bestandteile zerfallen, die einem und demselben Büschel angehören. Wählt man nun $\lambda_1 \lambda_2 \dots \lambda_n$ so, daß P_{n-4} irreduzibel ausfällt, so gehören F und P_{n-4} nicht demselben Büschel an, da ihre Ordnungen verschieden sind. Also ist die Schar irreduzibel.

Es seien $t_1 t_2 \dots t_{2n}$ ($n = 5, 6, 7$) $2n$ Doppeltangenten eines syzygetischen $2n$ -tupels, und Q_n sei eine durch die $4n$ Berührungspunkte gehende Kurve n -ter Ordnung. Dann definiert $\lambda Q_n^2 + t_1 t_2 \dots t_{2n} = 0$ ein Büschel von Kurven $2n$ -ter Ordnung. Es sei $\lambda Q_n^2 + t_1 t_2 \dots t_{2n} = 0$ eine noch durch einen Punkt von F gehende Kurve. Dann erhält man, genau wie in den Fällen $n = 2, 3, 4$, eine Identität

$$\lambda Q_n^2 + t_1 t_2 \dots t_{2n} = F P_{2n-4},$$

wobei P_{2n-4} eine Kurve $(2n - 4)$ -ter Ordnung ist. Jede t_i berührt die Kurve P_{2n-4} in $n - 2$ Punkten, und die Kurve Q_n geht noch durch die $n(n - 2)$ Berührungspunkte von P_{2n-4} mit t_i .

(Eingegangen am 13. 11. 1940.)

Unitärintvarianten im allgemeinen Euklidischen Raum.

Von

Hidegorô Nakano in Tokyo.

Kürzlich hat F. Wecken gezeigt, daß die bisherigen Formen von Unitärintvarianten selbstadjungierter Operatoren des Hilbertschen Raumes im nicht separablen Raum nicht mehr bestehen, und er hat eine neue Form von Unitärintvarianten im allgemeinen Euklidischen Raum durch Einführung von „Spektralbereichen“ aufgestellt ¹⁾.

Ich habe in einer früheren Abhandlung ²⁾ eine neue Form von Unitärintvarianten im Hilbertschen Raum angegeben: Man definiert zuerst die Dimension eines Rings von Projektionsoperatoren und die Dimensionalzerlegung eines Rings. N sei ein hypermaximaler normaler Operator mit dem Eigenring \mathfrak{R} ³⁾, der die Spektralschar von N ist. Entsprechend der Dimensionalzerlegung von \mathfrak{R} erhält man die Dimensionalzerlegung der komplexen Ebene G . Diese Zerlegung der komplexen Ebene ist die Unitärintvariante von N .

Diese Methode kann man nicht ohne Änderung im allgemeinen Euklidischen Raum anwenden, denn man benutzt für Dimensionalzerlegung eines Rings die Eigenschaft, daß der Eigenring eines hypermaximalen normalen Operators universal ⁴⁾ ist. Daher muß man im Falle des allgemeinen Euklidischen Raumes diese Methode etwas verändern, was der Zweck dieser Abhandlung ist.

Im folgenden bezeichnen wir mit \mathfrak{E} einen allgemeinen Euklidischen Raum, d. h. einen linearen Raum, der den Postulaten A, B und E von J. von Neumann ⁵⁾ genügt.

¹⁾ F. Wecken, Unitärintvarianten selbstadjungierter Operatoren. Math. Annalen **116** (1939), S. 422–455.

²⁾ H. Nakano, Unitärintvarianten hypermaximaler normaler Operatoren im Hilbertschen Raum. Diese Abhandlung soll in den Annals of Math. gedruckt werden. Im folgenden bezeichnen wir sie mit U. N. H.

³⁾ H. Nakano, Über Abelsche Ringe von Projektionsoperatoren. Proc. Phys.-Math. Soc. Japan **21** (1939), S. 357–375.

⁴⁾ Ein Ring von Projektionsoperatoren \mathfrak{U} heißt universal, wenn für jedes System — mit beliebiger Mächtigkeit — von Projektionsoperatoren $\{P_i\}$ in \mathfrak{U} die Summe $\sum P_i$ und das Produkt $\prod P_i$ beide auch zu \mathfrak{U} gehören. Siehe H. Nakano, Funktionen mehrerer hypermaximaler normaler Operatoren (im folgenden mit F. m. O. bezeichnet). Proc. Phys.-Math. Soc. Japan **21** (1939), S. 713–728.

⁵⁾ J. von Neumann, Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren. Math. Annalen **102** (1930), S. 49–131.

§ 1.

Dimensionalzerlegung universaler Ringe.

Zuerst betrachten wir nur universale Ringe. \mathfrak{U} sei ein universaler Ring von Projektionsoperatoren mit dem Maximaloperator M , d. h. $M = \sum_{P \in \mathfrak{U}} P$. Ein Projektionsoperator L heißt im folgenden mit \mathfrak{U} vertauschbar, wenn L mit jedem Projektionsoperator in \mathfrak{U} vertauschbar ist. Wie in der früheren Abhandlung⁶⁾ verstehen wir unter dem Nebenring $L\mathfrak{U}$ für einen mit \mathfrak{U} vertauschbaren Projektionsoperator L den Ring, der aus allen LR für beliebiges $R \in \mathfrak{U}$ besteht. Dann kann man leicht einsehen, daß jeder Nebenring eines universalen Rings auch universal ist.

Definition 1. Zwei Nebenringe $L_1\mathfrak{U}$ und $L_2\mathfrak{U}$ eines universalen Rings \mathfrak{U} heißen isomorph und wir schreiben

$$L_1\mathfrak{U} \sim L_2\mathfrak{U},$$

wenn für jedes $R \in \mathfrak{U}$ $L_1R = 0$ mit $L_2R = 0$ gleichbedeutend ist.

Definition 2. Eine Kardinalzahl α heißt die Minimaldimension eines universalen Rings \mathfrak{U} , wenn für jede Nichtlimites-Kardinalzahl $\beta \leq \alpha$ stets derartige β Nebenringe $\{L_i\mathfrak{U}\}$ existieren, daß für $i \neq j$ stets $L_iL_j = 0$, und für alle i $L_i\mathfrak{U} \sim \mathfrak{U}$ gilt, und wenn keine $\gamma (> \alpha)$ derartige Nebenringe existieren.

Definition 3. Ein universaler Ring \mathfrak{U} heißt gleichmäßig dimensional mit der Dimension α , wenn für jedes $R \in \mathfrak{U}$ die Minimaldimension des Nebenrings $R\mathfrak{U}$ stets α ist. Besonders nennen wir im Falle $\alpha = 1$ den universalen Ring \mathfrak{U} einfach.

Aus diesen Definitionen kann man ganz ähnlich wie in der früheren Abhandlung⁷⁾ die Dimensionalzerlegung jedes universalen Rings gewinnen. So ergibt sich

Satz 1. \mathfrak{U} sei ein universaler Ring mit dem Maximaloperator M . Man kann jeder Kardinalzahl α ein Element $R_\alpha \in \mathfrak{U}$ eindeutig derart zuordnen, daß der Nebenring $R_\alpha\mathfrak{U}$ gleichmäßig dimensional mit der Dimension α und für $\alpha \neq \beta$ stets $R_\alpha R_\beta = 0$, und $\sum_\alpha R_\alpha = M$ ist.

Satz 2. Dafür, daß ein universaler Ring \mathfrak{U} mit dem Maximaloperator M einfach ist, ist notwendig und hinreichend, daß für jeden mit \mathfrak{U} vertauschbaren Projektionsoperator L der Operator ML zu \mathfrak{U} gehört.

⁶⁾ H. Nakano, U. N. H., Definition 2.

⁷⁾ H. Nakano, U. N. H., Satz 5.

Beweis. Wenn ein universaler Ring \mathfrak{U} mit dem Maximaloperator M mit einem Projektionsoperator L vertauschbar ist, so gehören $M_1 = \sum_{\substack{RL=0 \\ R \in \mathfrak{U}}} R$ und $M_2 = \sum_{\substack{RL=M \\ R \in \mathfrak{U}}} R$ beide zu \mathfrak{U} , und es gilt

$$M - M_1 \geq ML \geq M_2,$$

$$(M - M_1 - M_2)\mathfrak{U} \sim L(M - M_1 - M_2)\mathfrak{U} \sim (M - ML)(M - M_1 - M_2)\mathfrak{U}.$$

Daher muß, wenn \mathfrak{U} einfach ist, $M - M_1 = M_2$ sein, folglich gehört $ML = M - M_1 = M_2$ zu \mathfrak{U} .

Wenn \mathfrak{U} nicht einfach ist, so gibt es ein derartiges $R \in \mathfrak{U}$, daß die Minimaldimension von $R\mathfrak{U}$ mindestens 2 ist. Dann gibt es einen derartigen mit \mathfrak{U} vertauschbaren Projektionsoperator L , daß

$$R\mathfrak{U} \sim MLR\mathfrak{U} \sim (M - ML)R\mathfrak{U}$$

ist. Daher gehört ML offenbar nicht zu \mathfrak{U} .

§ 2.

Separable Ringe.

Wir haben schon in der früheren Abhandlung⁸⁾ die Separabilität eines Rings definiert. Nun ist diese Definition für die vorliegende Untersuchung nicht bequem. Daher wollen wir eine andere geben, die mehr verlangt. Man kann aber die Identität beider Definitionen leicht beweisen, wenn man die Ungleichung $2^{\aleph_0} < 2^{\aleph_1}$ annimmt.

Definition 4. Ein universaler Ring \mathfrak{U} heißt separabel, wenn nicht mehr als abzählbar unendlich viele derartige Elemente $\{R_i\}$ in \mathfrak{U} existieren, so daß für $i \neq j$ stets $R_i R_j = 0$ ist.

Nach dieser Definition kann man leicht einsehen, daß jeder universale Unterring eines universalen separablen Rings auch separabel ist.

Satz 3. *Dafür, daß ein universaler Ring \mathfrak{U} separabel ist, ist notwendig und hinreichend, daß ein Element f in \mathfrak{E} existiert, so daß für jedes $R (\neq 0) \in \mathfrak{U}$ stets $Rf \neq 0$ gilt.*

Beweis. Wenn ein universaler Ring \mathfrak{U} separabel ist, so kann man wie im Hilfssatz 1 in der früheren Abhandlung⁹⁾ ein derartiges Element f in \mathfrak{E} finden, daß für jedes R in \mathfrak{U} $Rf \neq 0$ ist.

⁸⁾ H. Nakano, Über Abelsche Ringe ..., Satz 4.

⁹⁾ H. Nakano, U. N. H.

Umgekehrt, wenn ein solches Element f existiert, so muß \mathfrak{U} separabel sein, denn für jedes derartige System von $\{R_i\}$ in \mathfrak{U} , so daß für $i \neq j$ stets $R_i R_j = 0$ ist, gilt $\|f\|^2 \geq \sum_i \|R_i f\|^2$. Daher muß R_i bis auf höchstens abzählbar unendlich viele verschwinden.

Satz 4. Für jeden universalen Ring \mathfrak{U} gibt es mindestens ein Element $R_0 (\neq 0) \in \mathfrak{U}$, so daß der Nebenring $R_0 \mathfrak{U}$ separabel ist. Solches R_0 heißt ein *separables Element* von \mathfrak{U} .

Beweis. M sei der Maximaloperator von \mathfrak{U} , und f_0 sei ein derartiges Element in \mathfrak{E} , daß $M f_0 \neq 0$ ist. Alle Elemente $R \in \mathfrak{U}$, für die $R f_0 = 0$ ist, bilden auch einen universalen Ring \mathfrak{U}_0 mit dem Maximaloperator M_0 . Da \mathfrak{U} universal ist, stimmt \mathfrak{U}_0 mit dem Nebenring $M_0 \mathfrak{U}$ überein. Setzt man $R_0 = M - M_0$, so ist offenbar $\mathfrak{U} \ni R_0 \neq 0$, und nach Satz 3 ist der Nebenring $R_0 \mathfrak{U}$ separabel.

Aus diesem Satz ergibt sich sofort der folgende

Satz 5. \mathfrak{U} sei ein universaler Ring mit dem Maximaloperator M . Es gibt stets ein derartiges System von Elementen $\{R_i\}$ in \mathfrak{U} , daß für $i \neq j$ stets $R_i R_j = 0$ gilt, $R_i \mathfrak{U}$ alle separabel sind und $\sum_i R_i = M$ ist.

Definition 5. Ein universaler Ring \mathfrak{U} heißt *fundamental*, wenn eine derartige Folge von höchstens abzählbar unendlich vielen Elementen R_1, R_2, \dots in \mathfrak{U} existiert, so daß \mathfrak{U} der kleinste, R_1, R_2, \dots umfassende, universale Ring ist.

Satz 6. Wenn ein universaler Ring \mathfrak{U} fundamental ist, so ist für jedes f in \mathfrak{E} die aus allen Rf für $R \in \mathfrak{U}$ aufgespannte, abgeschlossene lineare Mannigfaltigkeit höchstens separabel, d. h. ein Hilbertscher oder endlich dimensionaler Raum.

Beweis. Da \mathfrak{U} fundamental ist, gibt es eine Folge von höchstens abzählbar unendlich vielen Elementen R_1, R_2, \dots in \mathfrak{U} , so daß \mathfrak{U} der kleinste, R_1, R_2, \dots umfassende, universale Ring ist. Bezeichnet man mit L_1, L_2, \dots die Projektionsoperatoren, welche aus allen Produkten von endlich vielen unter R_1, R_2, \dots bestehen, und mit Q den Projektionsoperator der aus $L_1 f, L_2 f, \dots$ aufgespannten, abgeschlossenen linearen Mannigfaltigkeit \mathfrak{M} , so gilt $Q R_i L_j f = R_i Q L_j f$ für alle i, j , und daraus folgt, daß die R_i alle mit Q vertauschbar sind. Nach Satz 3, 6 der früheren Abhandlung¹⁰⁾ erhält man dann, daß Q mit allen Elementen in \mathfrak{U} vertauschbar ist. Daher gilt für $R \in \mathfrak{U}$

$$Rf = R(R_1 + R_2 + \dots)f = RQ(R_1 + R_2 + \dots)f = QRf.$$

Folglich gehört Rf zu \mathfrak{M} , wo \mathfrak{M} offenbar höchstens separabel ist.

Satz 7. Dafür, daß ein fundamentaler universaler separabler Ring \mathfrak{U} einfach ist, ist notwendig und hinreichend, daß es ein Element f in \mathfrak{E} gibt, so

¹⁰⁾ H. Nakano, F. m. O.

daß der Maximaloperator M von \mathfrak{U} mit dem Projektionsoperator der aus allen Rf für $R \in \mathfrak{U}$ aufgespannten, abgeschlossenen linearen Mannigfaltigkeit übereinstimmt.

Beweis. Wenn ein universaler Ring \mathfrak{U} separabel ist, so gibt es nach Satz 3 ein derartiges Element f , daß $Rf \neq 0$ ist für jedes $R (\neq 0)$ in \mathfrak{U} . Der Projektionsoperator M_0 der aus allen Rf für $R \in \mathfrak{U}$ aufgespannten, abgeschlossenen linearen Mannigfaltigkeit ist dann mit \mathfrak{U} vertauschbar. Wenn \mathfrak{U} einfach ist, so muß M_0 nach Satz 2 zu \mathfrak{U} gehören. Folglich ist $M_0 = M$, da für alle $R (\neq 0) \in \mathfrak{U}$ ja $M_0 R \neq 0$ sein sollte.

Umgekehrt, wenn M für ein f in \mathfrak{E} der Projektionsoperator der aus allen Rf für $R \in \mathfrak{U}$ aufgespannten, abgeschlossenen linearen Mannigfaltigkeit ist, so ist diese abgeschlossene lineare Mannigfaltigkeit nach Satz 6 höchstens separabel, da \mathfrak{U} fundamental ist. In diesem Falle haben wir schon in der früheren Abhandlung¹¹⁾ bewiesen, daß \mathfrak{U} einfach ist.

Satz 8. Wenn ein fundamentaler, universaler, separabler Ring \mathfrak{U} gleichmäßig dimensional mit der Dimension α ist, so gibt es derartige α Nebenringe $\{M_i, \mathfrak{U}\}$, daß $M_i M_j = 0$ für $i \neq j$, $M_i \mathfrak{U} \sim \mathfrak{U}$, $M_i \mathfrak{U}$ einfach, und $\sum M_i$ der Maximaloperator von \mathfrak{U} ist. Hier ist M_i für ein passendes f_i in \mathfrak{E} der Projektionsoperator der aus allen Rf_i für $R \in \mathfrak{U}$ aufgespannten, höchstens separablen, abgeschlossenen linearen Mannigfaltigkeit.

Beweis. Im Falle $\alpha \leq \aleph_0$ haben wir die Behauptung schon in der früheren Abhandlung¹²⁾ bewiesen. Im allgemeinen Falle wollen wir ähnlich vorgehen.

Da \mathfrak{U} separabel ist, gibt es nach Satz 3 ein derartiges f_1 in \mathfrak{E} , daß für jedes $R (\neq 0) \in \mathfrak{U}$ $Rf_1 \neq 0$ ist. Bezeichnet man mit \mathfrak{M}_1 die aus allen Rf_1 für $R \in \mathfrak{U}$ aufgespannte, abgeschlossene lineare Mannigfaltigkeit, und mit M_1 den Projektionsoperator von \mathfrak{M}_1 , so ist \mathfrak{M}_1 höchstens separabel, und der Nebenring $M_1 \mathfrak{U}$ ist isomorph mit \mathfrak{U} und einfach. Wenn der Nebenring $(M - M_1)\mathfrak{U}$ mit \mathfrak{U} isomorph ist, so ist $(M - M_1)\mathfrak{U}$ auch separabel, und so gibt es ein derartiges Element f'_2 in \mathfrak{E} , daß $(M - M_1)Rf'_2 \neq 0$ für jedes $R (\neq 0) \in \mathfrak{U}$ gilt. Setzt man $f_2 = (M - M_1)f'_2$, und bezeichnet man mit \mathfrak{M}_2 die aus allen Rf_2 für $R \in \mathfrak{U}$ aufgespannte, abgeschlossene lineare Mannigfaltigkeit, und mit M_2 den Projektionsoperator von \mathfrak{M}_2 , so ist \mathfrak{M}_2 höchstens separabel. Der Nebenring $M_2 \mathfrak{U}$ ist isomorph mit \mathfrak{U} und einfach, und weiter gilt $M_1 M_2 = 0$. Ähnlich erhält man nach der transfiniten Induktion β_1 derartige Elemente $\{f_i\}$ und die entsprechenden abgeschlossenen linearen Mannigfaltigkeiten $\{\mathfrak{M}_i\}$ bzw. mit den Projektionsoperatoren $\{M_i\}$, daß \mathfrak{M}_i die aus allen Rf_i für $R \in \mathfrak{U}$ aufgespannte, abgeschlossene lineare Mannigfaltigkeit, der Nebenring $M_i \mathfrak{U}$

¹¹⁾ H. Nakano, U. N. H., Hilfssatz 3.

¹²⁾ H. Nakano, U. N. H., Hilfssatz 2, 4.

mit \mathfrak{U} isomorph und $M_i M_j = 0$ für $i \neq j$, und $(M - \sum_i M_i)\mathfrak{U}$ nicht mehr mit \mathfrak{U} isomorph ist. Nach Definition 2, 3 ist die Mächtigkeit β_1 nicht größer als α . Da \mathfrak{U} universal ist, bilden alle Elemente $R \in \mathfrak{U}$, die $(M - \sum_i M_i)R = 0$ genügen, einen universalen Ring mit dem Maximaloperator R_1 . R_1 gehört offenbar zu \mathfrak{U} , folglich ist der Nebenring $R_1 \mathfrak{U}$ gleichmäßig dimensional mit der Dimension α . Daher ist die Dimension der dem Projektionsoperator R_1 entsprechenden, abgeschlossenen linearen Mannigfaltigkeit nicht kleiner als α . Hier verstehen wir unter der Dimension einer abgeschlossenen linearen Mannigfaltigkeit die Mächtigkeit eines vollständigen Orthogonalsystems dieser Mannigfaltigkeit¹³⁾. Daher folgt aus $R_1 = \sum_i M_i R_1$, daß die Mächtigkeit β_1 nicht kleiner als α ist. Folglich stimmt die Mächtigkeit β_1 mit α überein.

Da der Nebenring $(M - R_1)\mathfrak{U}$ auch separabel und gleichmäßig dimensional mit der Dimension α ist, so gibt es wie oben ein derartiges $R_2 \in \mathfrak{U}$, daß $R_1 R_2 = 0$ gilt und α Elemente $\{f_i^{(2)}\}$ existieren, so daß für den Projektionsoperator $M_i^{(2)}$ der aus allen $R f_i$ für $R \in \mathfrak{U}$ aufgespannten, abgeschlossenen linearen Mannigfaltigkeit $\mathfrak{M}_i^{(2)}$

$$R_2 = \sum_i M_i^{(2)} R_2, \quad M_i^{(2)} M_j^{(2)} = 0 \quad (\text{für } i \neq j)$$

gilt.

Indem man dieses Verfahren fortsetzt, erhält man eine Folge von Elementen R_1, R_2, \dots in \mathfrak{U} , derart daß $R_i R_j = 0$ für $i \neq j$ und $M = R_1 + R_2 + \dots$ ist. Da \mathfrak{U} separabel ist, sind die R_1, R_2, \dots nicht mehr als abzählbar unendlich viele.

Setzt man

$$f_i^{(0)} = \frac{R_1 f_i}{2 \|f_i\|} + \frac{R_2 f_i^{(2)}}{2^2 \|f_i^{(2)}\|} + \dots,$$

und bezeichnet man mit $M_i^{(0)}$ den Projektionsoperator der aus allen $R f_i^{(0)}$ für $R \in \mathfrak{U}$ aufgespannten, abgeschlossenen linearen Mannigfaltigkeit $\mathfrak{M}_i^{(0)}$, so gilt

$$M_i^{(0)} = M_i R_1 + M_i^{(2)} R_2 + \dots,$$

folglich

$$M = \sum_i M_i^{(0)} \quad \text{und} \quad M_i^{(0)} M_j^{(0)} = 0 \quad \text{für } i \neq j.$$

Da $R f_i^{(0)} \neq 0$ für $R (\neq 0) \in \mathfrak{U}$ ist, so gilt $M_i^{(0)} \mathfrak{U} \sim \mathfrak{U}$. Damit haben wir den Satz bewiesen.

Nach obigem erhält man sofort den

Satz 9. Für jeden fundamentalen universalen Ring \mathfrak{U} gibt es mindestens einen Nebenring $L\mathfrak{U}$, der mit \mathfrak{U} isomorph und sogar einfach ist.

¹³⁾ F. Rellich, Spektraltheorie in nichtseparablen Räumen. Math. Annalen 110 (1935), S. 342–356.

§ 3.

Ideale von Borelmengen.

In dieser Abhandlung betrachten wir nur Borelmengen in der komplexen Ebene G . Man kann aber alle beschriebenen Tatsachen ohne weiteres nicht nur auf Borelmengen in einem abstrakten Raum, sondern auch auf Elemente in einem allgemeinen Booleschen Ring erweitern.

Definition 6. Ein System von Borelmengen p heißt ein **Ideal**, wenn p mit einer Borelmenge Z_0 auch alle Borelmengen $Z \supset Z_0$, und mit einer Folge höchstens abzählbar unendlich vieler Borelmengen Z_1, Z_2, \dots auch das Produkt $Z_1 Z_2 \dots$ enthält.

Definition 7. p_1 und p_2 seien zwei Ideale. Wenn alle Borelmengen in p_1 auch zu p_2 gehören, so sagt man, daß p_1 nicht kleiner als p_2 oder p_2 nicht größer als p_1 ist, und bezeichnet dies mit $p_1 \geq p_2$ oder $p_2 \leq p_1$.

Wenn $p_1 \geq p_2$ und sogar $p_1 \neq p_2$ ist, so sagt man, daß p_1 größer als p_2 oder p_2 kleiner als p_1 ist, und bezeichnet dies mit $p_1 > p_2$ oder $p_2 < p_1$.

Definition 8. Wenn ein Ideal p zwei fremde Borelmengen enthält, so enthält p die Nullmenge, und somit alle Borelmengen. Dieses Ideal p heißt **Nullideal** und man bezeichnet das Nullideal mit o .

Definition 9. Es sei eine Folge von Idealen p_1, p_2, \dots gegeben. Das Ideal, welches der Durchschnitt von p_1, p_2, \dots ist, heißt die **Summe** von p_1, p_2, \dots und man bezeichnet diese Summe mit $p_1 + p_2 + \dots$. Das Ideal, welches das p_1, p_2, \dots umfassende, kleinste Ideal ist, heißt das **Produkt** von p_1, p_2, \dots , und man bezeichnet dieses Produkt mit $p_1 p_2 \dots$.

Definition 10. Zwei Ideale p_1 und p_2 heißen **fremd**, wenn zwei derartige Borelmengen Z_1 und Z_2 existieren, so daß $Z_1 \in p_1, Z_2 \in p_2$, und $Z_1 Z_2 = 0$ ist, d. h. $p_1 p_2 = 0$.

Definition 11. Es sei ein Ideal p gegeben. Wenn für zwei Borelmengen Z_1 und Z_2 p eine derartige Borelmenge Z_0 enthält, daß $Z_1 Z_0 = Z_2 Z_0$ ist, so sagt man, daß die Borelmengen Z_1 und Z_2 miteinander **kongruent modulo p** sind, und bezeichnet dies mit

$$Z_1 \equiv Z_2 \text{ mod } p.$$

Nach Definition 11 ist $Z \in p$ offenbar mit

$$Z \equiv G \text{ mod } p$$

gleichbedeutend.

Satz 10. Es seien zwei Folgen höchstens abzählbar unendlich vieler Borelmengen Z'_1, Z'_2, \dots und Z''_1, Z''_2, \dots gegeben. Wenn für ein Ideal p

$$Z'_i \equiv Z''_i \text{ mod } p \quad (i = 1, 2, \dots)$$

gilt, so gilt auch

$$\begin{aligned} Z'_1 + Z'_2 + \dots &\equiv Z''_1 + Z''_1 + \dots \pmod{p}, \\ Z'_1 Z'_2 \dots &\equiv Z''_1 Z''_2 \dots \pmod{p}, \\ \varinjlim Z'_i &\equiv \varinjlim Z''_i \pmod{p}, \\ \varprojlim Z'_i &\equiv \varprojlim Z''_i \pmod{p}. \end{aligned}$$

Beweis. Wenn $Z'_i \equiv Z''_i \pmod{p}$ ist, so gibt es nach Definition 11 eine Borelmenge $Z'_i Z''_i \in p$, für die $Z'_i Z'_i = Z'_i Z''_i$ ist. Setzt man $Z_0 = Z'_1 Z'_2 \dots$, so gehört Z_0 auch zu p , und es gilt

$$Z'_i Z_0 = Z''_i Z_0, \quad (i = 1, 2, \dots);$$

daher auch

$$\begin{aligned} (Z'_1 + Z'_2 + \dots) Z_0 &= (Z''_1 + Z''_2 + \dots) Z_0, \\ (Z'_1 Z'_2 \dots) Z_0 &= (Z''_1 Z''_2 \dots) Z_0, \\ (\varinjlim Z'_i) Z_0 &= (\varinjlim Z''_i) Z_0, \\ (\varprojlim Z'_i) Z_0 &= (\varprojlim Z''_i) Z_0; \end{aligned}$$

damit haben wir die Behauptung bewiesen.

Satz 11. Wenn für $Z'_1 \supset Z'_2$, $Z''_1 \supset Z''_2$

$$Z'_1 \equiv Z''_1, \quad Z'_2 \equiv Z''_2 \pmod{p}$$

gilt, so gilt auch

$$Z'_1 - Z'_2 \equiv Z''_1 - Z''_2 \pmod{p}.$$

Satz 12. Wenn für höchstens abzählbar unendlich viele Ideale p_1, p_2, \dots zugleich

$$Z_1 \equiv Z_2 \pmod{p_i} \quad (i = 1, 2, \dots)$$

gilt, so ist auch

$$Z_1 \equiv Z_2 \pmod{p_1 + p_2 + \dots}.$$

Wenn für ein Ideal p

$$Z_1 \equiv Z_2 \pmod{p}$$

ist, so gilt auch für jedes Ideal $q \leq p$

$$Z_1 \equiv Z_2 \pmod{q}.$$

Beweis. Wenn $Z_1 \equiv Z_2 \pmod{p_i}$ ist, so gibt es eine Borelmenge $Z'_i \in p_i$, für die $Z_1 Z'_i = Z_2 Z'_i$ ist. Daher gilt auch

$$Z_1 (Z'_1 + Z'_2 + \dots) = Z_2 (Z'_1 + Z'_2 + \dots).$$

Da $Z'_1 + Z'_2 + \dots \in p_1 + p_2 + \dots$ sein sollte, so gilt

$$Z_1 \equiv Z_2 \pmod{p_1 + p_2 + \dots}.$$

Die letzte Behauptung folgt sofort aus Definition 11.

Satz 13. Wenn für ein Ideal p

$$Z_1 \equiv Z_2 \text{ und } Z_2 \equiv Z_3 \pmod{p}$$

zugleich gilt, so ist auch

$$Z_1 \equiv Z_3 \pmod{p}.$$

Beweis. Nach Definition 11 gibt es zwei Borelmengen Z' und Z'' in p , so daß

$$Z_1 Z' = Z_2 Z', \quad Z_2 Z'' = Z_3 Z''$$

ist. Daher gilt

$$Z_1 (Z' Z'') = Z_2 Z' Z'' = Z_3 (Z' Z'').$$

Da $Z' Z'' \in p$ sein sollte, erhält man $Z_1 \equiv Z_3 \pmod{p}$.

Satz 14. Es seien zwei Ideale p und q . Wenn $p \geq q$, $Z_1 \in q$, und $Z_1 \equiv Z_2 \pmod{p}$ ist, so gehört Z_2 auch zu q .

Beweis. Da $Z_1 \in q$ mit $Z_1 \equiv G \pmod{q}$ gleichbedeutend ist, so erhält man aus $Z_1 \equiv Z_2 \pmod{p}$ nach Satz 12

$$Z_1 \equiv Z_2 \pmod{q},$$

folglich

$$Z_2 \equiv G \pmod{q}.$$

Definition 12. Es sei p ein Ideal und Z_0 eine Borelmenge. Alle Borelmengen Z , die

$$ZZ_0 \equiv Z_0 \pmod{p}$$

genügen, bilden offenbar ein Ideal p_0 . Dieses Ideal nennen wir das **Nebenideal** p_0 von p zu der **primitiven** Borelmenge Z_0 , und Z_0 heißt eine **primitive** Borelmenge des Nebenideals p_0 über p .

Satz 15. Zwei Nebenideale p_{x_1} und p_{x_2} eines Ideals p zu den primitiven Borelmengen Z_1 bzw. Z_2 sind dann und nur dann fremd, wenn $Z_1 Z_2 \equiv 0 \pmod{p}$ ist.

Beweis. Wenn p_{x_1} und p_{x_2} fremd sind, so gibt es nach Definition 11 zwei derartige Borelmengen Z'_1 und Z'_2 , daß $Z'_1 \in p_{x_1}$, $Z'_2 \in p_{x_2}$, und $Z'_1 Z'_2 = 0$ ist. Da

$$Z'_1 Z_1 \equiv Z_1, \quad Z'_2 Z_2 \equiv Z_2 \pmod{p}$$

sein sollte, erhält man

$$Z_1 Z_2 \equiv Z'_1 Z'_2 Z_1 Z_2 \equiv 0 \pmod{p}.$$

Umgekehrt, wenn $Z_1 Z_2 \equiv 0 \pmod{p}$ ist, gilt nach Satz 12 $Z_1 Z_2 \equiv 0 \pmod{p_{x_1} p_{x_2}}$. Da $Z_1 Z_2 \in p_{x_1} p_{x_2}$ sein sollte, so muß $p_{x_1} p_{x_2} = 0$ sein, daher sind p_{x_1} und p_{x_2} fremd.

Satz 16. p sei ein Ideal, und p_{x_1}, p_{x_2}, \dots sei eine Folge von höchstens abzählbar unendlich vielen Nebenidealen von p (bzw. zu den primitiven Borelmengen Z_1, Z_2, \dots). $p_{x_1} + p_{x_2} + \dots$ und $p_{x_1} p_{x_2} \dots$ sind dann auch die Nebenideale von p zu den primitiven Borelmengen $Z_1 + Z_2 + \dots$ bzw. $Z_1 Z_2 \dots$.

Beweis. Aus $Z \equiv G \pmod{p_{x_1} + p_{x_2} + \dots}$ folgt nach Satz 12 $Z \equiv G \pmod{p_{x_i}}$ ($i = 1, 2, \dots$). Da nach Definition 12 $ZZ_i \equiv Z_i \pmod{p}$ ist, so gilt

$$Z(Z_1 + Z_2 + \dots) \equiv Z_1 + Z_2 + \dots \pmod{p}.$$

Da dieses Verfahren umkehrbar ist, erhält man die erste Behauptung.

Wenn $Z \in p_{x_1} p_{x_2} \dots$ ist, so gibt es nach Definition 9 eine Folge von Borelmengen Z'_1, Z'_2, \dots , derart, daß $Z'_i \in p_{x_i}$ ($i = 1, 2, \dots$) und $Z = Z'_1 Z'_2 \dots$ ist. Da $Z'_i Z_i \equiv Z_i \pmod{p}$ ist, so gilt auch

$$Z(Z_1 Z_2 \dots) \equiv (Z'_1 Z_1)(Z'_2 Z_2) \dots \equiv Z_1 Z_2 \dots \pmod{p}.$$

Umgekehrt, wenn diese Kongruenz besteht, so gibt es eine derartige Borelmenge $Z_0 \in p$, daß $Z(Z_1 Z_2 \dots) Z_0 = (Z_1 Z_2 \dots) Z_0$ ist. Daher gilt auch $Z(Z_1 Z_0)(Z_2 Z_0) \dots = (Z_1 Z_0)(Z_2 Z_0) \dots$. Da $Z_i Z_0 \equiv Z_i \pmod{p}$ ist, so gehört nach Satz 14 $Z_i Z_0$ zu p_{x_i} , und folglich $(Z_1 Z_0)(Z_2 Z_0) \dots$ zu $p_{x_1} p_{x_2} \dots$. Daher gehört Z nach Definition 6 auch zu p_{x_1}, p_{x_2}, \dots .

Satz 17. p_{x_1} und p_{x_2} seien zwei Nebenideale eines Ideals p zu den primitiven Borelmengen Z_1 bzw. Z_2 . Dafür, daß $p_{x_1} \leq p_{x_2}$ ist, ist notwendig und hinreichend, daß $Z_1 Z_2 \equiv Z_1 \pmod{p}$ ist. Im Falle $p_{x_1} \leq p_{x_2}$ gibt es ein einziges Nebenideal p_{x_3} von p , für das $p_{x_1} p_{x_2} = 0$ und $p_{x_1} + p_{x_2} = p_{x_3}$ ist. Dieses Ideal p_{x_3} bezeichnen wir mit $p_{x_2} - p_{x_1}$.

Beweis. Wenn $p_{x_1} \leq p_{x_2}$ ist, so gehört Z_2 zu p_{x_1} . Daher gilt $Z_2 Z_1 \equiv Z_1 \pmod{p}$. Umgekehrt wenn $Z_2 Z_1 \equiv Z_1 \pmod{p}$ ist, so gilt für jedes $Z \in p_{x_1}$,

$$ZZ_2 \equiv Z_2, ZZ_2 Z_1 \equiv ZZ_1 \pmod{p},$$

folglich

$$Z_1 \equiv Z_2 Z_1 \equiv ZZ_1 \pmod{p}.$$

Daher gehört Z zu p_{x_1} .

Im Falle $p_{x_1} \leq p_{x_2}$, wenn man $Z_3 = Z_2 - Z_2 Z_1$ setzt, so ist offenbar $p_{x_1} p_{x_2} = 0$ und $p_{x_3} \leq p_{x_2}$. Daher ist $p_{x_1} + p_{x_3} \leq p_{x_2}$. Andererseits gilt für jedes $Z \in p_{x_1} + p_{x_3}$,

$$ZZ_1 \equiv Z_1, ZZ_3 \equiv Z_3 \pmod{p},$$

folglich

$$Z(Z_3 + Z_1 Z_2) \equiv Z_3 + Z_1 Z_2 \pmod{p}.$$

und weiter

$$ZZ_2 \equiv Z_2 \pmod{p}, \text{ d. h. } Z \in p_{x_2}.$$

Daher ist $p_{x_2} = p_{x_1} + p_{x_3}$. Die Eindeutigkeit des Nebenideals p_{x_3} kann man aus Satz 15 und 10 leicht schließen.

§ 4.

Separable Ideale.

Definition 13. Ein Ideal p heißt *separabel*, wenn man für jedes System von Borelmengen $\{Z_i\}$ aus $Z_i Z_j \equiv 0 \pmod{p}$ (für $i \neq j$) stets schließen kann, daß $Z_i \equiv 0 \pmod{p}$ bis auf höchstens abzählbar unendlich viele Z_i gilt.

Satz 18. Wenn man für jedes System von Borelmengen $\{Z_i\}$ stets aus $Z_i Z_j = 0$ für $i \neq j$ schließen kann, daß $Z_i \equiv 0 \pmod{p}$ bis auf höchstens abzählbar unendlich viele Z_i gilt, so ist das Ideal p separabel.

Beweis. Es sei eine transfinit wohlgeordnete Folge von \aleph_1 Borelmengen Z_1, Z_2, \dots gegeben, die für $i \neq j$ stets $Z_i Z_j \equiv 0 \pmod{p}$, und für alle i $Z_i \not\equiv 0 \pmod{p}$ erfüllen. Da für jede Ordnungszahl i ($< \Omega$) die Mächtigkeit aller Ordnungszahlen j ($< i$) höchstens abzählbar unendlich ist, so erhält man

$$\left(\sum_{j < i} Z_j \right) Z_i \equiv 0 \pmod{p}.$$

Setzt man $Z'_i = Z_i - \left(\sum_{j < i} Z_j \right) Z_i$, so gilt

$$Z'_i \equiv Z_i \not\equiv 0 \pmod{p} \quad \text{und} \quad Z'_i Z'_j = 0 \quad \text{für} \quad i \neq j.$$

Daher muß, unter der Voraussetzung des Satzes, p separabel sein.

Aus diesem Satz folgt sofort der

Satz 19. Wenn ein Ideal p separabel ist, sind alle Ideale $q < p$ auch separabel.

Satz 20. Wenn ein Ideal p separabel ist, so ist jedes Ideal $q < p$ ein Nebenideal von p .

Beweis. Wenn man aus q nach der transfiniten Induktion eine derartige Folge von Borelmengen $Z_1 \supset Z_2 \supset \dots$ auswählt, daß $Z_i \in q$ und $Z_i - Z_{i+1} \not\equiv 0 \pmod{p}$ ist, so kann diese Folge Z_1, Z_2, \dots nicht mehr als abzählbar unendlich sein, da p separabel ist. Setzt man $Z_0 = Z_1 Z_2 \dots$, so ist Z_0 auch eine Borelmenge, $Z_0 \in q$ und für jedes $Z \in q$ ist $Z_0 Z \equiv Z_0 \pmod{p}$. Umgekehrt, wenn für eine Borelmenge Z $Z_0 Z \equiv Z_0 \pmod{p}$ ist, so gehört nach Satz 14 $Z_0 Z$ zu q , folglich Z nach Definition 6 auch zu q . Daher ist $q = p_{Z_0}$.

Satz 21. Für jede Folge von höchstens abzählbar unendlich vielen, separablen Idealen p_1, p_2, \dots ist die Summe $p_1 + p_2 + \dots$ auch separabel.

Beweis. $\{Z_i\}$ sei ein derartiges System von Borelmengen, daß für $i \neq j$ stets $Z_i Z_j = 0$ ist.

Da p_i separabel ist, gilt für ein bestimmtes i $Z_j \equiv 0 \pmod{p_i}$ bis auf höchstens abzählbar unendlich viele j . Daher ist

$$Z_j \equiv 0 \pmod{p_i} \quad (i = 1, 2, \dots)$$

bis auf höchstens abzählbar unendlich viele j . Folglich gilt nach Satz 12

$$Z_j \equiv 0 \bmod p_1 + p_2 + \dots$$

bis auf höchstens abzählbar unendlich viele j . Daher ist $p_1 + p_2 + \dots$ nach Satz 18 separabel.

§ 5.

Spektralsysteme von Maßoperatoren.

Nun betrachten wir einen Maßoperator $E(Z)$, wie wir ihn schon in der früheren Abhandlung¹⁴⁾ definiert haben, d. h. $E(Z)$ ist eine Schar von, jeder Borelmenge Z zugeordneten, Projektionsoperatoren, die den folgenden Bedingungen genügen:

- (1) $E(0) = 0,$
- (2) $E(Z_1 + Z_2 + \dots) = E(Z_1) + E(Z_2) + \dots$
- (3) $E(Z_1 Z_2 \dots) = E(Z_1) E(Z_2) \dots$

für höchstens abzählbar unendlich viele Borelmengen Z_1, Z_2, \dots

Für alle Borelmengen Z bildet die Schar der Projektionsoperatoren $E(Z)$ einen **Borelring**; d. h. einen, höchstens abzählbar unendlich viele passende Projektionsoperatoren umfassenden, kleinsten Ring von Projektionsoperatoren¹⁵⁾. Diesen Borelring bezeichnen wir mit \mathfrak{B} , und mit \mathfrak{U} den **U-Ring** von \mathfrak{B} , d. h. den \mathfrak{B} umfassenden, kleinsten universalen Ring von Projektionsoperatoren¹⁵⁾. Man kann leicht einsehen, daß der universale Ring \mathfrak{U} fundamental ist.

Wir nennen ein Produkt von irgendwelchen — mit beliebiger Mächtigkeit — Elementen \mathfrak{B} einen **δ -Operator** von $E(Z)$ oder \mathfrak{B} und bezeichnen das System aller δ -Operatoren mit \mathfrak{B}_δ . Dann ist offenbar $\mathfrak{B} \subset \mathfrak{B}_\delta \subset \mathfrak{U}$. Man kann auch leicht einsehen, daß jede Summe höchstens abzählbar unendlich vieler δ -Operatoren wieder ein δ -Operator ist. \mathfrak{B}_δ ist aber nicht immer ein Ring von Projektionsoperatoren.

Satz 22. *Jedes separable Element von \mathfrak{U} ist ein δ -Operator von \mathfrak{B} .*

Beweis. Da jedes Element von \mathfrak{U} als die Summe eines passenden Systems von δ -Operatoren sich darstellen läßt¹⁶⁾, so ist ein separables Element R von \mathfrak{U} die Summe eines Systems von δ -Operatoren $\{P_i\}$. Hier kann man annehmen, daß $\{P_i\}$ jede Summe von höchstens abzählbar unendlich vielen

¹⁴⁾ H. Nakano, Zur Eigenwerttheorie normaler Operatoren. Proc. Phys.-Math. Soc. Japan 21 (1939), S. 315–339.

¹⁵⁾ H. Nakano, F. m. O., S. 724. B -Ring. U -Ring.

¹⁶⁾ H. Nakano, F. m. O., Satz 13.

δ -Operatoren in $\{P_i\}$ auch enthält. Wenn man nach der transfiniten Induktion aus $\{P_i\}$ eine derartige transfinite Folge $P'_1 < P'_2 < \dots$ auswählt, daß für keinen δ -Operator $P'' \in \{P_i\}$ gleichzeitig für alle i $P'_i < P'' \leq R$ gilt, so kann diese Folge P'_1, P'_2, \dots nicht mehr als abzählbar unendlich sein, weil nach Voraussetzung R separabel ist. Setzt man $P_0 = P'_1 + P'_2 + \dots$, so gehört P_0 auch zu $\{P_i\}$ und für jedes i gilt $P_i \leq P_0$. Daher ist $P_0 = R$.

Definition 14. Für jeden δ -Operator P eines Maßoperators $E(Z)$ bilden alle Borelmengen Z , für die $PE(Z) = P$ gilt, offenbar ein Ideal \mathfrak{p} von Borelmengen. Dieses Ideal \mathfrak{p} heißt das Ideal von P , und P heißt der Normoperator des Ideals \mathfrak{p} über $E(Z)$, indem man P mit $N_{E(Z)}\mathfrak{p}$ bezeichnet.

Es gilt dann offenbar

$$N_{E(Z)}\mathfrak{p} = \prod_{Z \in \mathfrak{p}} E(Z).$$

Daher sind die Ideale von zwei verschiedenen δ -Operatoren auch voneinander verschieden.

Aus Definition 14 folgt sofort der

Satz 23. Für zwei δ -Operatoren P_1 und P_2 mit den Idealen \mathfrak{p}_1 bzw. \mathfrak{p}_2 ist die Ungleichung $\mathfrak{p}_1 < \mathfrak{p}_2$ mit $P_1 < P_2$ gleichbedeutend.

Satz 24. Es sei P ein δ -Operator mit dem Ideal \mathfrak{p} . Dafür, daß $Z_1 \equiv Z_2 \pmod{\mathfrak{p}}$ für zwei Borelmengen Z_1, Z_2 ist, ist notwendig und hinreichend, daß $E(Z_1)P = E(Z_2)P$ ist.

Beweis. Aus $Z_1 \equiv Z_2 \pmod{\mathfrak{p}}$ folgt nach Definition 11, daß $Z_1 Z_0 = Z_2 Z_0$ für eine passende Borelmenge $Z_0 \in \mathfrak{p}$ gilt. Daher gilt

$$E(Z_1)E(Z_0) = E(Z_1 Z_0) = E(Z_2 Z_0) = E(Z_2)E(Z_0).$$

Da $E(Z_0) \geq P$ sein sollte, erhält man $E(Z_1)P = E(Z_2)P$.

Umgekehrt, wenn $E(Z_1)P = E(Z_2)P$ ist, so gilt

$$E(Z_1)P = E(Z_1)E(Z_2)P = E(Z_1 Z_2)P,$$

folglich

$$E(Z_1 - Z_1 Z_2)P = 0.$$

Setzt man $Z_0 = G - (Z_1 - Z_1 Z_2)$, so ist Z_0 auch eine Borelmenge, und es gilt

$$E(Z_0)P = E(G)P - E(Z_1 - Z_1 Z_2)P = P.$$

Daher gehört Z_0 zu \mathfrak{p} . Da $Z_0(Z_1 - Z_1 Z_2) = 0$ ist, so gilt mithin nach Definition 11 $Z_1 Z_2 \equiv Z_1 \pmod{\mathfrak{p}}$. Ähnlich erhält man $Z_2 Z_1 \equiv Z_2 \pmod{\mathfrak{p}}$, und dann nach Satz 13 $Z_1 \equiv Z_2 \pmod{\mathfrak{p}}$.

Satz 25. Es sei eine Folge höchstens abzählbar unendlich vieler δ -Operatoren P_1, P_2, \dots von $E(Z)$ bzw. mit den Idealen $\mathfrak{p}_1, \mathfrak{p}_2, \dots$ gegeben. Dann wird das Ideal von $P_1 + P_2 + \dots$ durch $\mathfrak{p}_1 + \mathfrak{p}_2 + \dots$ gegeben.

Beweis. Wenn $Z_0 \equiv G \bmod p_1 + p_2 + \dots$ für eine Borelmenge Z_0 gilt, so ist für alle i $Z_0 \equiv G \bmod p_i$. Daher gilt nach Satz 24

$$E(Z_0) P_i = P_i \quad (i = 1, 2, \dots),$$

folglich

$$E(Z_0) (P_1 + P_2 + \dots) = P_1 + P_2 + \dots.$$

Umgekehrt folgt aus dieser Gleichung

$$E(Z_0) P_i = P_i \quad (i = 1, 2, \dots).$$

Daher gilt nach Satz 24

$$Z_0 \equiv G \bmod p_i \quad (i = 1, 2, \dots),$$

und dann nach Satz 12

$$Z_0 \equiv G \bmod p_1 + p_2 + \dots.$$

Damit ist der Satz bewiesen.

Satz 26. Es sei P ein δ -Operator mit dem Ideal p . Das Nebenideal p_Z von p zur primitiven Borelmenge Z_0 ist das Ideal von $PE(Z_0)$.

Beweis. Aus $Z \in p_{Z_0}$ folgt nach Definition 12, $ZZ_0 \equiv Z_0 \bmod p$. Daher gilt nach Satz 24 $E(ZZ_0)P = E(Z_0)P$, nämlich $E(Z)\{E(Z_0)P\} = \{E(Z_0)P\}$, und dieses Verfahren läßt sich umkehren.

Satz 27. Dafür, daß ein δ -Operator P mit dem Ideal p ein separables Element von \mathcal{U} ist, ist notwendig und hinreichend, daß p separabel ist.

Beweis. Wenn für ein System von Borelmengen $\{Z_i\}$ $Z_i Z_j \equiv 0 \bmod p$ ($i \neq j$) gilt, so ist nach Satz 24 $E(Z_i)E(Z_j)P = 0$, nämlich

$$\{E(Z_i)P\}\{E(Z_j)P\} = 0 \quad (i \neq j).$$

Daher, wenn $P\mathcal{U}$ separabel ist, gilt $E(Z_i)P = 0$ bis auf höchstens abzählbar viele i . Nach Satz 24 ist dann $Z_i \equiv 0 \bmod p$ bis auf höchstens abzählbar unendlich viele i . Daher ist p separabel.

Es sei $\{R_i\}$ ein derartiges System von Projektionsoperatoren in \mathcal{U} , daß für alle i $P \geq R_i > 0$, und für $i \neq j$ $R_i R_j = 0$ ist. Nach Satz 4 gibt es ein derartiges separables Element P_i von \mathcal{U} , daß $R_i \geq P_i > 0$ ist. Dann ist P_i nach Satz 22 ein δ -Operator von $E(Z)$. Bezeichnet man das Ideal von P_i mit p_i , so ist nach Satz 23 $p_i \leq p$. Wenn p separabel ist, so ist nach Satz 19 p_i ein Nebenideal von p . Z_i sei eine primitive Borelmenge von p_i über p . Dann ist nach Satz 26 $P_i = PE(Z_i)$. Aus $P_i P_j = 0$ ($i \neq j$) folgt für $i \neq j$

$$E(Z_i Z_j)P = E(Z_i)E(Z_j)P = P_i P_j = 0.$$

Daher gilt nach Satz 24

$$Z_i Z_j \equiv 0 \bmod p \quad (i \neq j).$$

Andererseits gilt, da nach Voraussetzung $P_i \neq 0$ sein sollte, $Z_i \neq 0 \bmod p$. Da p separabel ist, kann das System $\{Z_i\}$ nicht mehr als abzählbar unendlich sein. Daher ist $P\mathfrak{U}$ separabel.

Satz 28. *Es sei p_1, p_2, \dots eine Folge höchstens abzählbar unendlich vieler separabler Ideale. Wenn p_i das Ideal eines δ -Operators P_i des Maßoperators $E(Z)$ ist, so ist das Produkt $p_1 p_2 \dots$ auch das Ideal von $P_1 P_2 \dots$.*

Beweis. Nach Satz 25 ist $p_1 + p_2 + \dots$ das Ideal von $P_1 + P_2 + \dots$. Setzt man $p_0 = p_1 + p_2 + \dots$ und $P_0 = P_1 + P_2 + \dots$, so ist p_0 nach Satz 21 auch separabel. Daher ist p_i nach Satz 20 ein Nebenideal von p_0 . Bezeichnet man mit Z_i eine primitive Borelmenge von p_i über p_0 , so gilt nach Satz 24 $P_i = P_0 E(Z_i)$. Nach Satz 16 ist $p_1 p_2 \dots$ auch ein Nebenideal von p_0 zur primitiven Borelmenge $Z_1 Z_2 \dots$. Daher ist $p_1 p_2 \dots$ das Ideal von $PE(Z_1 Z_2 \dots) = \{PE(Z_1)\} \{PE(Z_2)\} \dots = P_1 P_2 \dots$.

Aus diesem Satz folgt sofort der

Satz 29. *Es seien P_1 und P_2 zwei separable Elemente von \mathfrak{U} mit den separablen Idealen p_1 bzw. p_2 . Dafür, daß $P_1 P_2 = 0$ ist, ist notwendig und hinreichend, daß zwei Ideale p_1 und p_2 zueinander fremd sind, d. h. $p_1 p_2 = 0$.*

Nach Satz 25, 28 erhält man leicht den

Satz 30. *Es seien P_1 und P_2 zwei separable Elemente von \mathfrak{U} mit den separablen Idealen p_1 bzw. p_2 . Wenn $P_1 \leq P_2$ ist, so ist $p_2 - p_1$ das Ideal von $P_2 - P_1$.*

Definition 15. Ein System separabler Ideale p heißt eine **Idealklasse**, wenn p mit einem Ideal $p \in p$ stets auch alle Ideale $q \leq p$, und mit einer Folge höchstens abzählbar unendlich vieler Ideale $p_1, p_2, \dots \in p$ stets auch die Summe $p_1 + p_2 + \dots$ enthält.

Definition 16. Ein System separabler Ideale $\{p_i\}$ in einer Idealklasse p heißt ein **Basissystem** von p , wenn für $i \neq j$ stets $p_i p_j = 0$ gilt und p die alle p_i umfassende, kleinste Idealklasse ist.

Satz 31. *Jede Idealklasse besitzt ein Basissystem.*

Beweis. $p_1 (\neq 0)$ sei ein Ideal in p . Alle Ideale p in p , für die $p_1 p = 0$ ist, bilden offenbar eine Idealklasse $p_1 \subset p$. $p_2 (\neq 0)$ sei ein Ideal in p_1 . Alle Ideale p in p_1 , für die $p_2 p = 0$ ist, bilden auch eine Idealklasse p_2 . Ähnlich fortsetzend, erhält man nach der transfiniten Induktion ein derartiges System $\{p_i\}$, daß für $i \neq j$ stets $p_i p_j = 0$ ist, und für ein Ideal p in p aus $p_i p = 0$ für alle i $p = 0$ folgt. p_0 sei ein beliebiges Ideal in p . Da p_0 separabel ist, gilt $p_i p_0 = 0$ bis auf höchstens abzählbar unendlich viele i . Es sei nun p_{i_1}, p_{i_2}, \dots die Folge aller der Ideale in $\{p_i\}$, für die $p_{i_l} p_0 \neq 0$ ($l = 1, 2, \dots$) ist. Dann gilt für alle i $\{p_0 - p_0(p_{i_1} + p_{i_2} + \dots)\} p_i = 0$, folglich muß

$$p_0 - p_0(p_{i_1} + p_{i_2} + \dots) = 0$$

sein, d. h.

$$p_0 \leq p_{i_1} + p_{i_2} + \dots$$

Daher ist $\{p_i\}$ ein Basissystem von p .

Definition 17. R sei ein Projektionsoperator in \mathfrak{U} . Alle separablen Ideale, deren jedes das Ideal eines separablen Elementes P von $R\mathfrak{U}$ ist, bilden offenbar eine Idealklasse. Diese Idealklasse nennen wir die **Idealklasse von R über den Maßoperator $E(Z)$** .

Satz 32. Für zwei Elemente R_1, R_2 in \mathfrak{U} ist $R_1 = R_2$ mit Übereinstimmung der Idealklassen von R_1 und R_2 gleichbedeutend.

Beweis. Wenn $R_1 - R_1 R_2 \neq 0$ ist, so gibt es nach Satz 4 ein separables Element P des Nebenrings $(R_1 - R_1 R_2)\mathfrak{U}$. Das separable Ideal p von P gehört dann zur Idealklasse von R_1 , aber nicht zur Idealklasse von R_2 nach Satz 29.

Definition 18. Die Dimensionalzerlegung des U-Rings \mathfrak{U} sei $R_1, R_2, \dots, R_\alpha, \dots$, und p_α sei die Idealklasse von R_α . $p_1, p_2, \dots, p_\alpha, \dots$ heißt das **Spektralsystem des Maßoperators $E(Z)$** .

§ 6.

Unitärinvarianten.

Es sei N ein hypermaximaler normaler Operator. Bezeichnet man mit $E(Z)$ den Maßoperator von N , so läßt N sich in der Form

$$(Nf, g) = \int_{\mathfrak{E}} z d(E(Z)f, g)$$

für jedes f im Definitionsbereich von N und für beliebiges g in \mathfrak{E} darstellen¹⁷⁾

Definition 19. Es sei N ein hypermaximaler normaler Operator mit dem Maßoperator $E(Z)$. Das Spektralsystem des Maßoperators $E(Z)$ heißt das **Spektralsystem von N** .

Dieses Spektralsystem eines hypermaximalen normalen Operators ist unitärinvariant, d. h.

Satz 33. Dafür, daß zwei hypermaximale normale Operatoren N_1 und N_2 unitär äquivalent sind, d. h. dafür daß für einen passenden unitären Operator U $N_1 = U^* N_2 U$ ist, ist notwendig und hinreichend, daß N_1 und N_2 dasselbe Spektralsystem haben.

Beweis. $E_1(Z)$ und $E_2(Z)$ seien die Maßoperatoren bzw. von N_1 und N_2 . Man kann leicht aus Eigenwertdarstellungen schließen: Dafür, daß für einen

¹⁷⁾ H. Nakano, Zur Eigenwerttheorie . . . , Satz 4.

unitären Operator $U N_1 = U^* N_2 U$ ist, ist notwendig und hinreichend, daß $E_1(Z) = U^* E_2(Z) U$ gilt für jede Borelmenge Z .

Für einen unitären Operator U sei $E_1(Z) = U^* E_2(Z) U$ für jede Borelmenge Z . \mathfrak{p} sei das separable Ideal eines beliebigen δ -Operators P_2 von $E_2(Z)$. Dann ist $P_2 = \prod_{Z \in \mathfrak{p}} E_2(Z)$. Setzt man $P_1 = \prod_{Z \in \mathfrak{p}} E_1(Z)$, so ist $\|P_1 f\|$ für jedes f in \mathfrak{E} die untere Grenze¹⁸⁾ von $\{\|E_1(Z) f\|\}$ für $Z \in \mathfrak{p}$, folglich die untere Grenze von $\|U^* E_2(Z) U f\| = \|E_2(Z) U f\|$ für $Z \in \mathfrak{p}$. Da diese untere Grenze andererseits mit $\|P_2 U f\|$ übereinstimmen muß, so erhält man für jedes f in \mathfrak{E}

$$\|P_1 f\| = \|P_2 U f\| = \|U^* P_2 U f\|.$$

Da $U^* P_2 U$ auch ein Projektionsoperator ist, folgt hieraus $P_1 = U^* P_2 U$.

Wenn für eine Borelmenge Z $P_1 \leq E_1(Z)$ ist, so gilt auch $U P_1 U^* \leq U E_1(Z) U^*$, d. h. $P_2 \leq E_2(Z)$. Daher gehört Z zu \mathfrak{p} . Dies besagt, daß \mathfrak{p} auch das Ideal des δ -Operators P_1 von $E_1(Z)$ ist.

Wenn der Nebenring $P_2 E_2(Z)$ gleichmäßig dimensional mit der Dimension α ist, so kann man aus $E_1(Z) = U^* E_2(Z) U$ leicht einsehen, daß der Nebenring $P_1 E_1(Z)$ auch gleichmäßig dimensional mit der Dimension α ist.

Umgekehrt gilt auch, daß jedes separable Ideal \mathfrak{p} über $E_1(Z)$ auch ein separables Ideal über $E_2(Z)$ ist, und im Falle, daß der Nebenring $\{N_{E_1(Z)} \mathfrak{p}\} E_1(Z)$ gleichmäßig dimensional mit der Dimension α ist, der Nebenring $\{N_{E_2(Z)} \mathfrak{p}\} E_2(Z)$ auch gleichmäßig dimensional mit der Dimension α ist. Daher haben $E_1(Z)$ und $E_2(Z)$ dasselbe Spektralsystem.

Nun nehmen wir an, daß $E_1(Z)$ und $E_2(Z)$ dasselbe Spektralsystem $p_1, p_2, \dots, p_\alpha, \dots$ haben. $\{p_i^\alpha\}$ sei ein Basissystem der Idealklasse p_α . Setzt man

$$P_1^{i, \alpha} = \prod_{Z \in \mathfrak{p}_i^\alpha} E_1(Z), \quad P_2^{i, \alpha} = \prod_{Z \in \mathfrak{p}_i^\alpha} E_2(Z),$$

so sind die Nebenringe $P_1^{i, \alpha} E_1(Z)$ und $P_2^{i, \alpha} E_2(Z)$ beide separabel und gleichmäßig dimensional mit der Dimension α . Nach Satz 8 gibt es dann derartige α Projektionsoperatoren $P_{1, k}^{i, \alpha}$ ($k = 1, 2, \dots$), daß

$$P_{1, k}^{i, \alpha} P_{1, l}^{i, \alpha} = 0 \quad (k \neq l),$$

$$P_{1, k}^{i, \alpha} E_1(Z) \sim P_1^{i, \alpha} E_1(Z), \quad P_1^{i, \alpha} = \sum_k P_{1, k}^{i, \alpha},$$

und der Nebenring $P_{1, k}^{i, \alpha} E_1(Z)$ einfach ist.

Über $P_2^{i, \alpha}$ gibt es ebenso derartige α Projektionsoperatoren $P_{2, k}^{i, \alpha}$ ($k = 1, 2, \dots$), daß

$$P_{2, k}^{i, \alpha} P_{2, l}^{i, \alpha} = 0 \quad (k \neq l),$$

$$P_{2, k}^{i, \alpha} E_2(Z) \sim P_2^{i, \alpha} E_2(Z), \quad P_2^{i, \alpha} = \sum_k P_{2, k}^{i, \alpha},$$

und der Nebenring $P_{2, k}^{i, \alpha} E_2(Z)$ einfach ist.

¹⁸⁾ H. Nakano, F. m. O., Satz 2.

Da beide Nebenringe $P_{1,k}^{\alpha} E_1(Z)$ und $P_{2,k}^{\alpha} E_2(Z)$ separabel und einfach sind und sogar dieselben Nullmengen haben, so gibt es zwei Elemente $f_{1,k}^{\alpha}$ und $f_{2,k}^{\alpha}$ in \mathfrak{E} , für die

$$P_{1,k}^{\alpha} f_{1,k}^{\alpha} = f_{1,k}^{\alpha}, \quad P_{2,k}^{\alpha} f_{2,k}^{\alpha} = f_{2,k}^{\alpha},$$

$$\|E_1(Z) f_{1,k}^{\alpha}\| = \|E_2(Z) f_{2,k}^{\alpha}\| \quad (\text{für alle } Z),$$

und $P_{1,k}^{\alpha}$, $P_{2,k}^{\alpha}$ die Projektionsoperatoren von, aus allen $E_1(Z) f_{2,k}^{\alpha}$ bzw. $E_2(Z) f_{1,k}^{\alpha}$ für beliebige Borelmengen Z aufgespannten, abgeschlossenen linearen Mannigfaltigkeiten sind, was wir schon in der früheren Abhandlung¹⁹⁾ bewiesen haben. Durch Zuordnung

$$E_1(Z) f_{1,k}^{\alpha} \leftrightarrow E_2(Z) f_{2,k}^{\alpha} \quad (\text{für alle } i, \alpha, k, Z)$$

erhält man einen unitären Operator U , für den

$$E_1(Z) f_{1,k}^{\alpha} = U^* E_2(Z) f_{2,k}^{\alpha} \quad (\text{für alle } i, \alpha, k, Z)$$

gilt. Hieraus folgt

$$E_1(Z_0) U^* E_2(Z) f_{2,k}^{\alpha} = E_1(Z_0) E_1(Z) f_{1,k}^{\alpha}$$

$$= E_1(Z_0 Z) f_{1,k}^{\alpha} = U^* E_2(Z_0 Z) f_{2,k}^{\alpha} = U^* E_2(Z_0) \{E_2(Z) f_{2,k}^{\alpha}\}$$

für beliebige Borelmengen Z_0 . Da alle Elemente $E_2(Z) f_{2,k}^{\alpha}$ für beliebige i, α, k, Z den ganzen Raum \mathfrak{E} aufspannen, gilt

$$E_1(Z_0) U^* f = U^* E_2(Z_0) f, \quad \text{d. h.} \quad E_1(Z_0) = U^* E_2(Z_0) U$$

für jede Borelmenge Z_0 , was zu beweisen war.

Es sei N_1, N_2, \dots eine Abelsche Folge hypermaximaler normaler Operatoren. In der früheren Abhandlung²⁰⁾ haben wir einen Maßoperator $E(M)$ auf dem komplexen Torusraum T von Elementen (z_1, z_2, \dots) konstruiert, derart daß die Operatoren N_i sich in der Form

$$(N_i f, g) = \int_T z_i d(E(M) f, g) \quad (i = 1, 2, \dots)$$

für jedes f im Definitionsbereich von N_i und für beliebiges g in \mathfrak{E} darstellen lassen. Wenn man in Betracht zieht, daß für jeden Kreis in T $C(C_1, C_2, \dots)$, wo C_i bis auf endlich viele mit der ganzen komplexen Ebene G übereinstimmen, und die anderen C_{j_1}, \dots, C_{j_k} die Kreise mit den Mittelpunkten $z_{j_1}^{(0)}, \dots, z_{j_k}^{(0)}$ und den Radien $r_{j_1}^{(0)}, \dots, r_{j_k}^{(0)}$ sind, $E(C)$ der Projektionsoperator

¹⁹⁾ H. Nakano, U. N. H., Satz 5.

²⁰⁾ H. Nakano, F. m. O.

Mathematische Annalen, 118.

der abgeschlossenen linearen Mannigfaltigkeit ist, die aus allen für N_{j_1}, \dots, N_{j_k} gleichzeitig totalsinnvollen,

$$\|(N_{j_l} - z_{j_l}^{(0)})^n f\| < r_{j_l}^{(0)n} \|f\| \quad \left(\begin{array}{l} n = 1, 2, \dots \\ l = 1, 2, \dots, k \end{array} \right)$$

genügenden Elementen f besteht, so kann man wie im Falle eines Operators²¹⁾ leicht beweisen, daß der Maßoperator $E(M)$ eindeutig bestimmt wird.

Das Spektralsystem des Maßoperators $E(M)$ heißt das Spektralsystem der Abelschen Folge hypermaximaler normaler Operatoren N_1, N_2, \dots .

Satz 34. Es seien N_1, N_2, \dots und N'_1, N'_2, \dots zwei Abelsche Folgen hypermaximaler normaler Operatoren. Dafür, daß für einen passenden unitären Operator U gleichzeitig

$$N'_i = U^* N_i U \quad (i = 1, 2, \dots)$$

ist, ist notwendig und hinreichend, daß die beiden Folgen N_1, N_2, \dots und N'_1, N'_2, \dots dasselbe Spektralsystem haben.

Beweis. Die Maßoperatoren bzw. von N_1, N_2, \dots und N'_1, N'_2, \dots seien $E(M)$ und $E'(M)$. Nach der Eigenwertdarstellung und der Eindeutigkeit des Maßoperators kann man leicht einsehen: Dafür, daß für einen unitären Operator U gleichzeitig $N'_i = U^* N_i U$ ist, ist es notwendig und hinreichend, daß $E'(M) = U^* E(M) U$ für jede Borelmenge M in T gilt. Daher kann man ganz ähnlich wie Satz 33 diesen Satz beweisen.

§ 7.

Reguläre Ideale.

Wir haben schon gezeigt, daß durch jeden Maßoperator $E(Z)$ ein Spektralsystem $p_1, p_2, \dots, p_\alpha, \dots$ eindeutig bestimmt wird, wobei p_α eine der Kardinalzahl α zugeordnete Idealklasse ist. Nun ergibt sich die Frage: Welches System von jeder Kardinalzahl zugeordneten Idealklassen kann das Spektralsystem eines Maßoperators $E(Z)$ in einem passenden allgemeinen Euklidischen Raum sein? Hier wollen wir diese Frage untersuchen.

Definition 20. Ein Ideal p heißt regulär, wenn es eine derartige beschränkte, nicht negative, totaladditive Mengenfunktion von Borelmengen $F(Z)$ gibt, daß

$$Z \equiv 0 \pmod{p}$$

mit $F(Z) = 0$ gleichbedeutend ist. Eine solche Mengenfunktion $F(Z)$ nennen wir ein Maß zum Ideal p .

²¹⁾ H. Nakano, Zur Eigenwerttheorie ..., S. 331–332.

Aus dieser Definition folgt sofort der

Satz 35. *Jedes reguläre Ideal ist separabel.*

Satz 36. *Dafür, daß ein Ideal p regulär ist, ist notwendig und hinreichend, daß es einen derartigen Maßoperator $E(Z)$ gibt, für den*

$$Z \equiv 0 \pmod{p}$$

mit $E(Z) = 0$ gleichbedeutend ist.

Beweis. $F(Z)$ sei ein Maß zum Ideal p , und

$$Z = G = Z_{i_1}, Z_{i_1 i_2}, Z_{i_1 i_2 i_3}, \dots, Z_{i_1 i_2 \dots i_n}, \dots \quad (i_n = 0, 1)$$

sei ein derartiges dyadisches Schema abzählbar unendlich vieler Borelmengen, daß die diese Borelmengen umfassende, kleinste totaladditive Mengenklaasse auch alle Borelmengen enthält, und

$$Z_{i_1 \dots i_n 1} Z_{i_1 \dots i_n 0} = 0, \quad Z_{i_1 \dots i_n} \supset Z_{i_1 \dots i_n i_{n+1}},$$

$$Z_{i_1 \dots i_n} = Z_{i_1 \dots i_n 0} + Z_{i_1 \dots i_n 1} \quad (i_n = 0, 1)$$

gilt. \mathfrak{H} sei ein Hilbertscher, nämlich separabler, Raum, und f sei ein Element in \mathfrak{H} mit der Norm $\sqrt{F(Z)}$. Da $F(Z) = F(Z_0) + F(Z_1)$ sein soll, so gibt es zwei zueinander orthogonale Elemente f_0 und f_1 mit der Norm $\sqrt{F(Z_0)}$ bzw. $\sqrt{F(Z_1)}$, für die $f = f_0 + f_1$ ist. Nun zerlege man den Raum \mathfrak{H} in zwei zueinander orthogonale separable Räume \mathfrak{H}_0 und \mathfrak{H}_1 , die f_0 bzw. f_1 enthalten.

Indem man dieses Verfahren fortsetzt, erhält man zwei dyadische Schemata von Elementen und separablen Räumen

$$\begin{aligned} f, f_{i_1}, f_{i_1 i_2}, \dots, f_{i_1 \dots i_n}, \dots \\ \mathfrak{H}, \mathfrak{H}_{i_1}, \mathfrak{H}_{i_1 i_2}, \dots, \mathfrak{H}_{i_1 \dots i_n}, \dots \end{aligned} \quad (i_n = 0, 1),$$

derart, daß $\mathfrak{H}_{i_1 \dots i_n 0}$ und $\mathfrak{H}_{i_1 \dots i_n 1}$ zueinander orthogonal sind, $\mathfrak{H}_{i_1 \dots i_n} = \mathfrak{H}_{i_1 \dots i_n 0} + \mathfrak{H}_{i_1 \dots i_n 1}$, $f_{i_1 \dots i_n} \in \mathfrak{H}_{i_1 \dots i_n}$, und $\|f_{i_1 \dots i_n}\| = \sqrt{F(Z_{i_1 \dots i_n})}$ ist. Man kann leicht einsehen, daß es einen und nur einen derartigen Maßoperator $E(Z)$ gibt, daß $E(Z_{i_1 \dots i_n})$ der Projektionsoperator des Raumes $\mathfrak{H}_{i_1 \dots i_n}$ im Raum \mathfrak{H} ist. Für diesen Maßoperator $E(Z)$ gilt offenbar $f_{i_1 \dots i_n} = E(Z_{i_1 \dots i_n})f$. Daher gilt auch $F(Z_{i_1 \dots i_n}) = \|E(Z_{i_1 \dots i_n})f\|^2$. Da $\|E(Z)f\|^2$ eine nicht negative, beschränkte, totaladditive Mengenfunktion ist, und für $Z_{i_1 \dots i_n}$ mit $F(Z)$ denselben Wert annimmt, so stimmt $\|E(Z)f\|^2$ mit $F(Z)$ für jede Borelmenge überein. p_0 sei das Ideal von $E(G)$ über $E(Z)$. Dann gilt offenbar $p \leq p_0$. Andererseits ist der Ring $\{E(Z)\}$ für alle Borelmengen universal und separabel. Daher ist p nach Satz 20 ein Nebenideal von p_0 zu einer primitiven Borelmenge Z_0 , d. h. $Z \equiv 0 \pmod{p}$ ist mit $E(Z_0)E(Z) = 0$ gleichbedeutend.

Umgekehrt, wenn p ein Ideal eines Maßoperators $E(Z)$ in einem separablen Raum ist, so ist nach Satz 3 p offenbar ein reguläres Ideal.

Satz 37. Für ein reguläres Ideal p_0 sind alle Ideale $p \leq p_0$ auch regulär.

Beweis. Da p_0 nach Satz 35 separabel sein muß, ist p nach Satz 20 ein Nebenideal von p_0 . Z_0 sei eine primitive Borelmenge von p über p_0 , und $F_0(Z)$ sei ein Maß zum Ideal p_0 . Setzt man

$$F(Z) = F_0(Z) - F_0(Z_0 Z),$$

so kann man leicht einsehen, daß $F(Z)$ ein Maß zum Ideal p ist.

Satz 38. Für eine Folge höchstens abzählbar unendlich vieler regulärer Ideale p_1, p_2, \dots ist $p_1 \dot{+} p_2 \dot{+} \dots$ auch regulär.

Beweis. $F_i(Z)$ sei ein Maß zum Ideal p_i . a_1, a_2, \dots sei eine derartige Folge positiver Zahlen, daß $\sum_{i=1}^{\infty} a_i F_i(G)$ konvergiert. Setzt man $F(Z) = a_1 F_1(Z) + a_2 F_2(Z) + \dots$, so kann man leicht einsehen, daß $F(Z)$ ein Maß zum Ideal $p_1 \dot{+} p_2 \dot{+} \dots$ ist.

Satz 39. Es sei ein System von jeder Kardinalzahl zugeordneten Idealklassen $p_1, p_2, \dots, p_\alpha, \dots$. Dafür, daß $p_1, p_2, \dots, p_\alpha, \dots$ das Spektralsystem eines hypermaximalen normalen Operators N in einem passenden allgemeinen Euklidischen Raum ist, ist notwendig und hinreichend, daß für $\alpha \neq \beta$ jedes Ideal in p_α zu allen in p_β fremd ist und daß die Ideale in p_α ($\alpha = 1, 2, \dots$) alle regulär sind.

Beweis. Diese Bedingung ist offenbar notwendig. Daß sie hinreichend ist, wollen wir beweisen. $\{p_\alpha\}$ sei ein Basissystem von p_α . Nach Satz 36 kann man derartige Hilbertsche Räume $\mathfrak{H}_{\alpha i}$ und Maßoperatoren $E_{\alpha i}(Z)$ in $\mathfrak{H}_{\alpha i}$ konstruieren, daß $p_{\alpha i}$ das Ideal von $E_{\alpha i}(Z)$ ist. Wenn man dabei je zwei $\mathfrak{H}_{\alpha i}$ zueinander orthogonal sein läßt und $\mathfrak{E} = \sum_{\alpha, i} \mathfrak{H}_{\alpha i}$ setzt, so hat der Maßoperator

$$E(Z) = \sum_{\alpha, i} E_{\alpha i}^h(Z) E_{\alpha i}$$

für die Projektionsoperatoren $E_{\alpha i}$ von $\mathfrak{H}_{\alpha i}$ in \mathfrak{E} $p_1, p_2, \dots, p_\alpha, \dots$ zum Spektralsystem. Da $E(G) = 1$ ist, erhält man einen hypermaximalen normalen Operator N eindeutig durch

$$(Nf, g) = \int_G zd(E(Z)f, g).$$

Das Spektralsystem dieses Operators N ist dann offenbar $p_1, p_2, \dots, p_\alpha, \dots$.

Zuletzt bemerken wir, daß nicht jedes separable Ideal regulär ist. Wir konstruieren ein nicht reguläres, aber separables Ideal.

Durchläuft I alle Borelmengen erster Kategorie, so bilden die Komplementärmenge $G - I$ zusammen ein Ideal p_I . Dieses Ideal p_I ist separabel. Denn für jedes System von Borelmengen $\{Z_i\}$, bei dem für $i \neq j$ $Z_i Z_j \equiv 0 \pmod{p_I}$ ist, gibt es ein solches System offener Punktmenge $\{G_i\}$, daß $Z_i \equiv G_i \pmod{p_I}$ ist, da man jede Borelmenge Z_i für eine passende offene Punktmenge G_i und für passende Borelmengen erster Kategorie I'_i, I''_i in der Form

$$Z_i = G_i - I'_i + I''_i$$

schreiben kann. Daher gilt $G_i G_j \equiv 0 \pmod{p_I}$ für $i \neq j$. Da eine offene Punktmenge nicht von erster Kategorie sein kann, gilt mithin $G_i G_j = 0$. Folglich gilt $G_i = 0$, d. h. $Z_i \equiv 0 \pmod{p_I}$, bis auf höchstens abzählbar unendlich viele i .

$F(Z)$ sei eine derartige nicht negative, beschränkte, totaladditive Mengenfunktion, daß $F(I) = 0$ ist für jede Borelmenge erster Kategorie I . Dann verschwindet $F(Z)$ offenbar für einen Punkt. Daher kann man wie die Cantorsche Punktmenge eine derartige perfekte, nirgend dichte, dyadische Punktmenge Z_0 konstruieren, daß $F(G) - F(Z_0)$ beliebig klein wird. Da Z_0 von erster Kategorie ist, so muß $F(G) = 0$ sein. Daher ist das Ideal p_I nicht regulär.

(Eingegangen am 2. 5. 1940.)

Über Approximation meromorpher Funktionen durch rationale Funktionen¹⁾.

Von

Ernst Lammel in Prag.

$\{a_\mu\}$, ($\mu = 1, 2, \dots$) sei eine Folge von Stellen aus $|z| \leq \varrho < R$ und $\{b_\mu\}$, ($\mu = 1, 2, \dots$) eine Folge von Stellen aus $|z| < R$, deren sämtliche Häufungspunkte auf $|z| = R$ liegen.

Jeder in $|z| < R$ meromorphen Funktion $f(z)$, deren sämtliche in $|z| < R$ gelegenen Pole unter den Stellen der Folge $\{b_\mu\}$, ($\mu = 1, 2, \dots$) auftreten sollen²⁾, lassen sich die rationalen Funktionen

$$(1) \quad s_n(z) = \begin{cases} C_0 & \text{für } n = 0, \\ C_0 + \sum_{\nu=1}^n C_\nu \prod_{\mu=1}^{\nu} \frac{z - a_\mu}{z - b_\mu} & \text{für } n \geq 1 \end{cases}$$

durch folgende Festsetzung eindeutig zuordnen: Tritt a_1 unter den Stellen $\{a_\mu\}$, ($\mu = 1, 2, \dots, n+1$) nur einmal auf, so soll $s_n(a_1) = f(a_1)$ sein. Kommt es dagegen κ -mal vor, so soll $s_n(z)$ mit $f(z)$ an der Stelle a_1 außerdem noch in den Ableitungen bis zur $(\kappa - 1)$ -ten Ordnung übereinstimmen.

Die rationalen Funktionen $s_n(z)$ sind die Partialsummen der Reihe

$$(2) \quad C_0 + \sum_{\nu=1}^{\infty} C_\nu \prod_{\mu=1}^{\nu} \frac{z - a_\mu}{z - b_\mu}.$$

Zu jeder in $|z| < R$ meromorphen Funktion $f(z)$, welche der Klasse $H\{b\}$ angehört, gibt es also eine und nur eine Reihe (2).

Wir fragen nach notwendigen und zugleich hinreichenden Bedingungen, welchen die Folge $\{a_\mu\}$, ($\mu = 1, 2, \dots$) aus $|z| \leq \varrho < R$ im Verein mit der Folge $\{b_\mu\}$, ($\mu = 1, 2, \dots$) aus $|z| < R$, für die $\lim_{\mu \rightarrow \infty} |b_\mu| = R$ ist, genügen muß, damit die einer jeden in $|z| < R$ meromorphen Funktion $f(z)$ der Klasse $H\{b\}$ eindeutig zugeordnete Reihe (2) für jeden Wert von z aus $|z| < R$,

¹⁾ Über Approximation meromorpher Funktionen durch rationale vgl. insbesondere das 8. Kapitel des Werkes von J. L. Walsh, *Interpolation and Approximation by Rational Functions in the Complex Domain* (American Mathematical Society Colloquium Publications, Volume XX), New York City, 1935.

²⁾ Die Gesamtheit dieser Funktionen wollen wir im folgenden als die Klasse $H\{b\}$ der in $|z| < R$ meromorphen Funktionen $f(z)$ bezeichnen.

welcher kein Punkt der Folge $\{b_\mu\}$, ($\mu = 1, 2, \dots$) ist, gegen $f(z)$ konvergiert. Wir werden zeigen, daß hierfür

$$(3) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\mu=1}^n (a_\mu^k - b_\mu^k) = 0; \quad k = 1, 2, \dots$$

eine sowohl notwendige als auch hinreichende Bedingung ist.

Es ist keine Beschränkung der Allgemeinheit, wenn wir, wie es in Hinkunft stets getan wird, $R = 1$ setzen.

Ist allgemeiner \mathfrak{B} ein beschränkter offener Bereich, dessen Berandung von einer geschlossenen Jordankurve C gebildet wird, und sind $\{a_\mu\}$, ($\mu = 1, 2, \dots$) und $\{b_\mu\}$, ($\mu = 1, 2, \dots$) zwei Folgen aus \mathfrak{B} , wobei von $\{a_\mu\}$, ($\mu = 1, 2, \dots$) kein Häufungspunkt, während von $\{b_\mu\}$, ($\mu = 1, 2, \dots$) sämtliche Häufungspunkte auf C liegen, so konvergiert dann und nur dann die einer jeden in \mathfrak{B} meromorphen Funktion $f(z)$ der Klasse $H\{b\}$ eindeutig zugeordnete Reihe (2) für jeden Wert von z aus \mathfrak{B} , welcher kein Punkt der Folge $\{b_\mu\}$, ($\mu = 1, 2, \dots$) ist, gegen $f(z)$, wenn bei konformer Abbildung $z = z(\zeta)$ von \mathfrak{B} auf den Einheitskreis $|\zeta| < 1$ die beiden Folgen $\{a_\mu\}$, ($\mu = 1, 2, \dots$) und $\{b_\mu\}$, ($\mu = 1, 2, \dots$) in Bildfolgen übergehen, welche den Bedingungen (3) genügen.

§ 1.

Integraldarstellung für die Entwicklungskoeffizienten $\{C_\nu\}$, ($\nu = 0, 1, 2, \dots$) und das Restglied.

Um für die Koeffizienten $\{C_\nu\}$, ($\nu = 0, 1, 2, \dots$) der zu einer in \mathfrak{B} meromorphen Funktion $f(z)$ der Klasse $H\{b\}$ gehörigen Reihe (2) und ihr Restglied eine Integraldarstellung zu erhalten, kann man nach R. Lagrange^{*)} so vorgehen: Man multipliziere die Identität

$$(4) \quad \frac{1}{\zeta - z} = \frac{1}{\zeta - a_1} + \frac{a_2 - b_1}{(\zeta - a_1)(\zeta - a_2)} \frac{z - a_1}{z - b_1} + \frac{a_3 - b_2}{(\zeta - a_2)(\zeta - a_3)} \frac{\zeta - b_1}{\zeta - a_1} \frac{z - a_1}{z - b_1} \frac{z - a_2}{z - b_2} +$$

$$+ \dots + \frac{a_{n+1} - b_n}{(\zeta - a_n)(\zeta - a_{n+1})} \prod_{\mu=1}^{n-1} \frac{\zeta - b_\mu}{\zeta - a_\mu} \prod_{\mu=1}^n \frac{z - a_\mu}{z - b_\mu} +$$

$$+ \frac{1}{\zeta - z} \frac{z - a_{n+1}}{\zeta - a_{n+1}} \prod_{\mu=1}^n \frac{z - a_\mu}{z - b_\mu} \frac{\zeta - b_\mu}{\zeta - a_\mu}$$

mit $\frac{1}{2\pi i} f(\zeta)$ und integriere dann längs einer in \mathfrak{B} verlaufenden rektifizierbaren geschlossenen Jordankurve L , welche die Folge $\{a_\mu\}$, ($\mu = 1, 2, \dots$) in ihrem

^{*)} Siehe R. Lagrange, Mémoire sur les séries d'interpolation. Acta Mathematica 64 (1935), S. 6.

Innern und die Folge $\{b_\mu\}$, ($\mu = 1, 2, \dots$) in ihrem Äußern enthält. Es ergibt sich so

$$f(z) = C_0 + \sum_{\nu=1}^n C_\nu \prod_{\mu=1}^{\nu} \frac{z-a_\mu}{z-b_\mu} + R_n(z),$$

wenn

$$(5a) \quad \begin{aligned} C_0 &= \frac{1}{2\pi i} \oint_L \frac{f(\zeta)}{\zeta - a_1} d\zeta, & C_1 &= \frac{a_2 - b_1}{2\pi i} \oint_L \frac{f(\zeta)}{(\zeta - a_1)(\zeta - a_2)} d\zeta, \\ C_\nu &= \frac{a_{\nu+1} - b_\nu}{2\pi i} \oint_L \frac{f(\zeta)}{(\zeta - a_1)(\zeta - a_{\nu+1}) \prod_{\mu=1}^{\nu-1} \frac{\zeta - a_\mu}{\zeta - b_\mu}} d\zeta; \quad \nu = 2, 3, \dots \end{aligned}$$

und

$$(5b) \quad R_n(z) = \frac{1}{2\pi i} \oint_L \frac{f(\zeta)}{\zeta - z} \frac{z - a_{n+1}}{\zeta - a_{n+1}} \prod_{\mu=1}^n \frac{\frac{z - a_\mu}{z - b_\mu}}{\frac{\zeta - a_\mu}{\zeta - b_\mu}} d\zeta$$

gesetzt wird.

Insbesondere folgt aus (5a, b) für die Funktion $f(z) = z$ die Entwicklung

$$(6) \quad \begin{aligned} z &= a_1 + (a_2 - b_1) \frac{z - a_1}{z - b_1} + (a_3 - b_2) \frac{z - a_1}{z - b_1} \frac{z - a_2}{z - b_2} + \dots \\ &\dots + (a_n - b_{n-1}) \prod_{\mu=1}^{n-1} \frac{z - a_\mu}{z - b_\mu} + (z - a_{n+1}) \prod_{\mu=1}^n \frac{z - a_\mu}{z - b_\mu}. \end{aligned}$$

§ 2.

Der Fall des Einheitskreises.

1. Die Bedingungen (3) sind notwendig.

Insbesondere gehört die Funktion $f(z) = z$ zur Klasse $H\{b\}$. Soll sie sich in eine Reihe (2) entwickeln lassen, so muß wegen (6) für jede Stelle z aus $|z| < 1$, die nicht in der Folge $\{b_\mu\}$, ($\mu = 1, 2, \dots$) auftritt,

$$\lim_{n \rightarrow \infty} \left| (z - a_{n+1}) \prod_{\mu=1}^n \frac{z - a_\mu}{z - b_\mu} \right| = 0$$

sein, woraus

$$(7) \quad \lim_{n \rightarrow \infty} \left| \sqrt[n]{\prod_{\mu=1}^n \frac{z - a_\mu}{z - b_\mu}} \right| \leq 1$$

folgt. Nun ist wegen $\lim_{\mu \rightarrow \infty} |b_\mu| = 1$ für jeden Wert von z aus $|z| < 1$

$$\lim_{\mu \rightarrow \infty} \left| \frac{z - b_\mu}{1 - \bar{b}_\mu z} \right| = 1$$

und deshalb

$$(8) \quad \lim_{n \rightarrow \infty} \sqrt[n]{\prod_{\mu=1}^n \frac{z - b_{\mu}}{1 - \bar{b}_{\mu} z}} = 1.$$

Mithin ergibt sich aus (7)

$$(9) \quad \lim_{n \rightarrow \infty} \left| \sqrt[n]{\prod_{\mu=1}^n \frac{z - a_{\mu}}{1 - \bar{b}_{\mu} z}} \right| \leq 1,$$

und zwar gilt diese Limesbeziehung, wie man leicht einsieht, für jeden Wert von z aus $|z| < 1$.

Jetzt zeigen wir, daß aus (9) für jeden Wert von z aus $|z| < 1$

$$(10) \quad \lim_{n \rightarrow \infty} \left| \sqrt[n]{\prod_{\mu=1}^n \frac{1 - \bar{a}_{\mu} z}{1 - \bar{b}_{\mu} z}} \right| \leq 1$$

erschlossen werden kann.

Wir setzen zur Abkürzung

$$\varphi_n(z) = \sqrt[n]{\prod_{\mu=1}^n \frac{1 - \bar{a}_{\mu} z}{1 - \bar{b}_{\mu} z}}; \quad n = 1, 2, \dots,$$

wobei immer der Funktionszweig gewählt werden soll, für welchen $\varphi_n(0) = 1$ ist.

Angenommen, für $z = z_1$ sei $\lim_{n \rightarrow \infty} |\varphi_n(z_1)| = 1 + g$, $g > 0$. Dann läßt sich aus der Funktionenfolge $\{\varphi_n(z)\}$, ($n = 1, 2, \dots$) eine auf jeder abgeschlossenen Punktmenge aus $|z| < 1$ gleichmäßig gegen eine in $|z| < 1$ reguläre Funktion $\varphi(z)$ konvergente Teilfolge $\{\varphi_{n_{\lambda}}(z)\}$, ($\lambda = 1, 2, \dots$) aussondern, so daß gleichzeitig

$$\lim_{\lambda \rightarrow \infty} |\varphi_{n_{\lambda}}(z_1)| = 1 + g$$

ist. Wegen des Maximumprinzips regulärer Funktionen ist

$$(11) \quad \max_{|z|=r} |\varphi(z)| \geq 1 + g, \text{ wenn } |z_1| < r < 1.$$

Da

$$|\varphi_n(z)| = \sqrt[n]{\prod_{\mu=1}^n \frac{z - a_{\mu}}{1 - \bar{b}_{\mu} z} \frac{1 - \bar{a}_{\mu} z}{1 - \bar{b}_{\mu} z}}$$

ist, so gilt

$$(12) \quad \lim_{n \rightarrow \infty} |\varphi_n(z)| \leq \frac{\lim_{n \rightarrow \infty} \sqrt[n]{\prod_{\mu=1}^n \frac{z - a_{\mu}}{1 - \bar{b}_{\mu} z}}}{\lim_{n \rightarrow \infty} \sqrt[n]{\prod_{\mu=1}^n \frac{z - a_{\mu}}{1 - \bar{a}_{\mu} z}}}.$$

Weil die Folge $\{a_\mu\}$, ($\mu = 1, 2, \dots$) in $|z| \leq \varrho < 1$ enthalten ist und für $\varrho \leq |z| < 1$ die Ungleichung

$$\left| \frac{z - a_\mu}{1 - \bar{a}_\mu z} \right| \geq \frac{|z| - \varrho}{1 - \varrho |z|}$$

besteht, so ergibt sich aus (12) an jeder Stelle z aus dem Kreisringe $\varrho \leq |z| < 1$ die Abschätzung

$$\lim_{n \rightarrow \infty} \left| \sqrt[n]{\prod_{\mu=1}^n \frac{z - a_\mu}{1 - \bar{a}_\mu z}} \right| \geq \frac{|z| - \varrho}{1 - \varrho |z|} \lim_{n \rightarrow \infty} |\varphi_n(z)|.$$

Die rechte Seite dieser Ungleichung ist aber wegen (11) bei passender Wahl von z sicher größer als $1 + \frac{g}{2}$, was in Widerspruch zu (9) steht.

Wir beweisen nun, daß aus (10)

$$(13) \quad \lim_{n \rightarrow \infty} |\varphi_n(z)| \geq 1$$

folgt und mithin für jede Stelle z aus $|z| < 1$

$$(14) \quad \lim_{n \rightarrow \infty} |\varphi_n(z)| = 1$$

ist.

Angenommen, für $z = z_1$ sei

$$\lim_{n \rightarrow \infty} |\varphi_n(z_1)| = 1 - g, \quad g > 0.$$

Dann läßt sich aus der Folge $\{\varphi_n(z)\}$, ($n = 1, 2, \dots$) eine Teilfolge $\{\varphi_{n_1}(z)\}$, ($\lambda = 1, 2, \dots$) aussondern, welche auf jeder abgeschlossenen Punktmenge aus $|z| < 1$ gleichmäßig gegen eine daselbst reguläre Grenzfunktion $\varphi(z)$ konvergiert und für die $|\varphi(z_1)| = 1 - g$ ist. Da $\varphi(0) = 1$, so muß es wegen des Maximumprinzips bei regulären Funktionen in $|z| < 1$ mindestens eine Stelle z_2 geben, an welcher $|\varphi(z_2)| > 1$ ist. Um so mehr müßte im Widerspruch zu (10)

$$\lim_{n \rightarrow \infty} |\varphi_n(z_2)| > 1$$

gelten.

log $\varphi_n(z)$ ist eine in $|z| < 1$ reguläre Funktion, die wir dadurch auch noch eindeutig machen, daß wir den Zweig nehmen, für welchen $\log \varphi_n(0) = 0$ ist. log $|\varphi_n(z)|$ ist als Realteil von log $\varphi_n(z)$ eine in $|z| < 1$ reguläre Potentialfunktion. Da die Folge $\{\log |\varphi_n(z)|\}$, ($n = 1, 2, \dots$) wegen (14) an jeder Stelle z aus $|z| < 1$ gegen Null konvergiert und auf jeder abgeschlossenen Menge aus $|z| < 1$ gleichmäßig beschränkt bleibt, so konvergiert die Folge $\{\log |\varphi_n(z)|\}$, ($n = 1, 2, \dots$) auf jeder abgeschlossenen Punktmenge aus $|z| < 1$ auch gleichmäßig gegen Null. Die Folge der konjugierten Potentialfunktionen, welche durch die Imaginärteile von log $\varphi_n(z)$ gegeben ist, kon-

vergiert mithin gleichmäßig auf jeder abgeschlossenen Punktmenge aus $|z| < 1$ gegen die gleiche Konstante, die wegen $\log \varphi_n(0) = 0$ den Wert Null haben muß.

Wir erhalten also schließlich

$$(15) \quad \lim_{n \rightarrow \infty} \varphi_n(z) = 1,$$

und zwar gleichmäßig auf $|z| \leq r < 1$.

Für das Bestehen von (15) sind aber die Bedingungen (3) sowohl notwendig als auch hinreichend⁴⁾.

2. Die Bedingungen (3) sind hinreichend.

Es soll nun gezeigt werden, daß die zu jeder Funktion $f(z)$ der Klasse $H\{b\}$ gehörige Reihe (2), wenn die Bedingungen (3) erfüllt sind, an jeder Stelle z aus $|z| < 1$, welche nicht in der Folge $\{b_\mu\}$, ($\mu = 1, 2, \dots$) auftritt, gegen $f(z)$ konvergiert.

$z = z_1$ sei eine solche Stelle. Wir wählen m zunächst so groß, daß für $\mu > m$ die Ungleichung $|b_\mu| > \text{Max}(|z_1|, \varrho)$ erfüllt ist. Dann verhält sich die Funktion

$$g_m(z) = \left\{ f(z) - \left(C_0 + \sum_{\nu=1}^{m-1} C_\nu \prod_{\mu=1}^{\nu} \frac{z - a_\mu}{z - b_\mu} \right) \right\} \prod_{\mu=1}^m \frac{z - b_\mu}{z - a_\mu}$$

in einem Kreise $|z| < r_m < 1$ regulär, welcher die Folge $\{a_\mu\}$, ($\mu = 1, 2, \dots$) und die Stelle $z = z_1$ als inneren Punkt enthält. $\{C_\nu\}$, ($\nu = 0, 1, 2, \dots$) ist dabei die Folge der Koeffizienten, welche in der $f(z)$ eindeutig zugeordneten Reihe (2) auftreten.

$g_m(z)$ besitzt die Entwicklung

$$g_m(z) = C_m + C_{m+1} \frac{z - a_{m+1}}{z - b_{m+1}} + \dots + C_n \prod_{\mu=m+1}^n \frac{z - a_\mu}{z - b_\mu} + R_{n-m}(z),$$

worin

$$R_{n-m}(z) = \frac{1}{2\pi i} \oint_L \frac{g_m(z)}{\zeta - z} \frac{z - a_{n+1}}{\zeta - a_{n+1}} \prod_{\mu=m+1}^n \frac{z - a_\mu}{\zeta - a_\mu} d\zeta$$

gesetzt wurde. Als Integrationsweg L wählen wir einen Kreis $|z| = l_m$, welcher die Stellen $z_1, a_1, a_2, \dots, a_{n+1}$ umschließt und in $|z| < r_m$ enthalten ist.

⁴⁾ Es läßt sich dies analog beweisen wie eine ähnliche Behauptung in E. Lammle, Über gewisse Reihen, welche die Potenzreihen als Grenzfall enthalten. Math. Zeitschr. 42 (1937), S. 394.

Wegen

$$\frac{z - a_\mu}{z - b_\mu} = \frac{z - \bar{a}_\mu}{1 - \bar{a}_\mu z} \cdot \frac{1 - \bar{b}_\mu z}{z - b_\mu} \cdot \frac{1 - \bar{a}_\mu z}{1 - \bar{b}_\mu z}, \quad \left| \frac{z - a_\mu}{1 - \bar{a}_\mu z} \right| \leq \frac{|z| + \varrho}{1 + \varrho|z|},$$

$$\frac{\zeta - a_\mu}{\zeta - b_\mu} = \frac{\zeta - \bar{a}_\mu}{1 - \bar{a}_\mu \zeta} \cdot \frac{1 - \bar{b}_\mu \zeta}{\zeta - b_\mu} \cdot \frac{1 - \bar{a}_\mu \zeta}{1 - \bar{b}_\mu \zeta}, \quad \left| \frac{\zeta - a_\mu}{1 - \bar{a}_\mu \zeta} \right| \leq \frac{|\zeta| + \varrho}{1 + \varrho|\zeta|},$$

wenn $\varrho < |\zeta| < 1$ ist,

und mit Rücksicht auf (8)⁵⁾ und (15) erhalten wir für das Restglied $R_{n-m}(z)$ von $g_m(z)$ an der Stelle $z = z_1$ die Abschätzung

$$\lim_{n \rightarrow \infty} \sqrt[n-m]{|R_{n-m}(z_1)|} \leq \frac{|z_1| + \varrho}{1 + \varrho|z_1|} \cdot \frac{l_m - \varrho}{1 - \varrho l_m}.$$

Ist hierin noch nicht

$$\frac{|z_1| + \varrho}{1 + \varrho|z_1|} < \frac{l_m - \varrho}{1 - \varrho l_m},$$

so kann dies sicher durch Vergrößerung von m erreicht werden. Damit ist dann unsere Behauptung bewiesen.

Aus dem Beweisgange ersieht man leicht, daß die Konvergenz der Reihe (2) auf jeder abgeschlossenen Punktmenge aus $|z| < 1$, welche die Punkte der Folge $\{b_\mu\}$, ($\mu = 1, 2, \dots$) weder enthält noch zu Häufungspunkten hat, sogar gleichmäßig ist.

§ 3.

Der Fall des einfach zusammenhängenden Jordanbereiches \mathfrak{B} .

1. Die angegebenen Bedingungen sind hinreichend.

Zunächst sprechen wir folgenden Hilfssatz aus, welchen wir später brauchen werden: Es seien $\{\alpha_\nu\}$, ($\nu = 1, 2, \dots$) und $\{\beta_\nu\}$, ($\nu = 1, 2, \dots$) zwei Folgen, die den Bedingungen

$$(16) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\mu=1}^n (\alpha_\mu^k - \beta_\mu^k) = 0; \quad k = 1, 2, \dots$$

genügen. Dann ist für jede in $|w| < 1$ reguläre und auf $|w| \leq 1$ stetige Funktion $f(w)$

$$(17) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\mu=1}^n \{f(\alpha_\mu) - f(\beta_\mu)\} = 0.$$

Die Voraussetzung (16) sagt aus, daß der Satz für Polynome richtig ist. Nun läßt sich aber jede in $|w| < 1$ reguläre und auf $|w| \leq 1$ stetige Funktion $f(w)$ auf $|w| \leq 1$ gleichmäßig durch Polynome approximieren⁶⁾.

⁵⁾ Man überzeugt sich leicht, daß in (8) die Konvergenz auf jeder abgeschlossene n Punktmenge aus $|z| < 1$ gleichmäßig ist.

⁶⁾ Vgl. das in Fußnote ¹⁾ zitierte Werk, S. 36.

Ist z eine feste Stelle aus \mathfrak{B} und $w(z, \zeta)$ eine derjenigen Abbildungsfunktionen, welche \mathfrak{B} so auf $|w| < 1$ abbilden, daß $\zeta = z$ in $w = 0$ übergeht, verhält sich $\frac{z - \zeta}{w(z, \zeta)}$ in $|w| < 1$ regulär und auf $|w| \leq 1$ stetig, weil bekanntlich $w(z, \zeta)$ eine in \mathfrak{B} reguläre und auf $\mathfrak{B} + C$ stetige Funktion ist. Auch $\log \frac{z - \zeta}{w(z, \zeta)}$ ist in $|w| < 1$ regulär und auf $|w| \leq 1$ stetig.

Sind die beiden Folgen $\{\alpha_\nu\}$, ($\nu = 1, 2, \dots$) und $\{\beta_\nu\}$, ($\nu = 1, 2, \dots$) aus $|w| < 1$ die Bildfolgen von $\{a_\nu\}$, ($\nu = 1, 2, \dots$) und $\{b_\nu\}$, ($\nu = 1, 2, \dots$) aus \mathfrak{B} bei der Abbildung $w = w(z, \zeta)$ und setzen wir voraus, daß die Bildfolgen den Bedingungen (16) genügen, so gilt nach unserem Hilfssatz

$$(18) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\mu=1}^n \left\{ \log \left| \frac{z - a_\mu}{w(z, a_\mu)} \right| - \log \left| \frac{z - b_\mu}{w(z, b_\mu)} \right| \right\} = 0,$$

weil

$$\log \left| \frac{z - \zeta}{w(z, \zeta)} \right| = \Re \left\{ \log \frac{z - \zeta}{w(z, \zeta)} \right\}$$

ist.

Setzen wir zur Abkürzung

$$\log |\varphi_n(z)| = \frac{1}{n} \sum_{\mu=1}^n \log \left| \frac{z - a_\mu}{w(z, a_\mu)} \frac{w(z, b_\mu)}{z - b_\mu} \right|; \quad n = 1, 2, \dots,$$

so bleibt die Folge von in \mathfrak{B} regulären Potentialfunktionen $\{\log |\varphi_n(z)|\}$, ($n = 1, 2, \dots$) auf jeder abgeschlossenen Punktmenge aus \mathfrak{B} gleichmäßig beschränkt und konvergiert nach (18) an jeder Stelle z aus \mathfrak{B} gegen Null. Hieraus folgt aber, daß auf jeder abgeschlossenen Punktmenge aus \mathfrak{B} die Konvergenz gegen Null sogar gleichmäßig ist. Mithin gilt in \mathfrak{B}

$$(19) \quad \lim_{n \rightarrow \infty} |\varphi_n(z)| = 1,$$

und zwar gleichmäßig auf jeder abgeschlossenen Punktmenge aus \mathfrak{B} .

Um nun zu beweisen, daß die so erhaltene Bedingung (19) tatsächlich hinreichend dafür ist, daß sich jede in \mathfrak{B} meromorphe Funktion der Klasse $H\{b\}$ in eine Reihe (2) entwickeln läßt, kann man analog wie im Falle des Einheitskreises vorgehen.

2. Die als hinreichend erkannten Bedingungen sind auch notwendig.

Soll sich *jede* in \mathfrak{B} meromorphe Funktion der Klasse $H\{b\}$ in eine Reihe (2) entwickeln lassen, so muß diese Eigenschaft insbesondere die Funktion $f(z) = \frac{1}{t-z}$ besitzen, worin t eine beliebige Stelle aus \mathfrak{B}' , dem Äußern von C , bedeutet. Nach der Identität von R. Lagrange (4) hat $f(z) = \frac{1}{t-z}$ eine Entwicklung, deren Restglied

$$R_n(z) = \frac{1}{t-z} \frac{z - a_{n+1}}{t - a_{n+1}} \prod_{\mu=1}^n \frac{z - a_\mu}{z - b_\mu} \frac{t - b_\mu}{t - a_\mu}$$

lautet. Soll also $f(z) = \frac{1}{t-z}$ in eine Reihe (2) entwickelbar sein, so muß deshalb für jede Stelle z aus \mathfrak{B} , welche nicht in der Folge $\{b_\mu\}$, ($\mu = 1, 2, \dots$) vorkommt,

$$\lim_{n \rightarrow \infty} \left| \frac{1}{t-z} \frac{z-a_{n+1}}{t-a_{n+1}} \prod_{\mu=1}^n \frac{z-a_\mu}{z-b_\mu} \frac{t-b_n}{t-a_n} \right| = 0$$

sein, woraus

$$(20) \quad \lim_{n \rightarrow \infty} \left| \sqrt[n]{\prod_{\mu=1}^n \frac{z-a_\mu}{z-b_\mu}} \right| \leq \lim_{n \rightarrow \infty} \left| \sqrt[n]{\prod_{\mu=1}^n \frac{t-a_\mu}{t-b_\mu}} \right|$$

folgt.

Wir wollen nun zunächst zeigen, daß wegen (20)

$$(21) \quad \lim_{n \rightarrow \infty} \left| \sqrt[n]{\prod_{\mu=1}^n \frac{t-a_\mu}{t-b_\mu}} \right| \geq 1$$

ist.

$w = w(z, a)$ sei eine der Funktionen, welche den Bereich \mathfrak{B} so auf den Einheitskreis $|w| < 1$ abbilden, daß der fest gewählte innere Punkt a aus \mathfrak{B} in $w = 0$ übergeht. Die Bildfolgen von $\{a_\nu\}$, ($\nu = 1, 2, \dots$) und $\{b_\nu\}$, ($\nu = 1, 2, \dots$) seien $\{\alpha_\nu\}$, ($\nu = 1, 2, \dots$) und $\{\beta_\nu\}$, ($\nu = 1, 2, \dots$). Da von der Folge $\{\alpha_\nu\}$, ($\nu = 1, 2, \dots$) kein Häufungspunkt und von $\{\beta_\nu\}$, ($\nu = 1, 2, \dots$) sämtliche Häufungspunkte auf $|w| = 1$ liegen, so läßt sich eine Folge von Kreisingen $\{K_\nu^{(1,2)}\}$, ($\nu = 1, 2, \dots$) finden, welche durch $r_\nu^{(1)} < |w| < r_\nu^{(2)}$, $r_\nu^{(2)} < r_{\nu+1}^{(1)}$ und $\lim_{\nu \rightarrow \infty} r_\nu^{(1)} = \lim_{\nu \rightarrow \infty} r_\nu^{(2)} = 1$ gegeben ist und wobei weder eine der α -noch der β -Stellen in einem der Kreisinge $K_\nu^{(1,2)}$ liegt. Wir führen noch die Folge der Kreise $|w| = \varrho_\nu$; $\nu = 1, 2, \dots$ mit $\varrho_\nu = \frac{1}{2}(r_\nu^{(1)} + r_\nu^{(2)})$ ein. Die Bereiche in \mathfrak{B} , deren Bildbereiche die Kreisinge $K_\nu^{(1,2)}$ sind, sollen mit $\mathfrak{B}_\nu^{(1,2)}$ bezeichnet werden und die Originalkurven von $|w| = \varrho_\nu$ mit C_ν .

Zur Abkürzung setzen wir

$$\Psi_n(z) = \sqrt[n]{\prod_{\mu=1}^n \frac{z-a_\mu}{z-b_\mu}}; \quad n = 1, 2, \dots$$

Die Funktionen $\Psi_n(z)$ sind in jedem der Bereiche $\mathfrak{B}_k^{(1,2)}$ regulär. Die Funktionenfolge $\{\Psi_n(z)\}$, ($n = 1, 2, \dots$) können wir dadurch in einem herausgegriffenen Bereiche $\mathfrak{B}_k^{(1,2)}$ zu einer Folge eindeutiger Funktionen machen, daß wir von jeder der Funktionen $\Psi(z)$ an einer Stelle aus $\mathfrak{B}_k^{(1,2)}$ z. B. den Zweig wählen, welcher daselbst den Hauptwert von $\Psi_n(z)$ als Funktionswert besitzt. Da die Funktionenfolge $\{\Psi_n(z)\}$, ($n = 1, 2, \dots$) auf jeder abgeschlossenen Punktmenge aus $\mathfrak{B}_k^{(1,2)}$ gleichmäßig beschränkt bleibt, so läßt sich eine

Teilfolge $\{\mathcal{P}_{n_1(k)}(z)\}$, ($\lambda = 1, 2, \dots$) aussondern, welche auf jeder abgeschlossenen Punktmenge aus $\mathfrak{B}_k^{(1,2)}$ gleichmäßig gegen eine in $\mathfrak{B}_k^{(1,2)}$ reguläre Grenzfunktion $\mathcal{P}^{(k)}(z)$ konvergiert.

In $\mathfrak{B}_k^{(1,2)}$ verläuft die Kurve C_k . \mathfrak{B}'_k sei das Äußere von C_k .

Auf \mathfrak{B}'_k üben wir zunächst die Transformation $t = \frac{1}{\tau}$ aus, und hierauf führen wir den so erhaltenen Bereich mit Hilfe von $\tau = \tau_k(w)$, wobei $\tau_k(0) = 0$ und $\tau'_k(0) > 0$ sein soll, in $|w| < 1$ über. Für $\frac{1}{\tau_k(w)}$ schreiben wir $t_k(w)$.

Alle Stellen der Folge $\{b_\mu\}$, ($\mu = 1, 2, \dots$), deren Zeiger $n \geq n_k$ ist, liegen in \mathfrak{B}'_k .

Es ist dann

$$\frac{1}{2\pi} \int_0^{2\pi} \log |\mathcal{P}_n [t_k(e^{i\theta})]| d\theta = \frac{1}{n} \sum_{\mu=n_k}^n \log |w_\mu^{(k)}|, \quad (7)$$

und da die Folge $\{\mathcal{P}_{n_1(k)}(z)\}$, ($\lambda = 1, 2, \dots$) auf C_k gleichmäßig konvergiert:

$$\frac{1}{2\pi} \int_0^{2\pi} \log |\mathcal{P}^{(k)} [t_k(e^{i\theta})]| d\theta = \lim_{\lambda \rightarrow \infty} \frac{1}{n_1(k)} \sum_{\mu=n_k}^{n_1(k)} \log |w_\mu^{(k)}|,$$

wenn $w_\mu^{(k)}$ der Bildpunkt ist, in welchen b_μ bei der Abbildung von \mathfrak{B}'_k auf $|w| < 1$ durch $t = t_k(w)$ übergeht.

Vermöge der Abbildung $t = t_k(w)$ entspricht der Kurve C eine in $|w| < 1$ verlaufende Kurve $C^{(k)}$ als Bildkurve. Ist

$$m_k = \min_{w \in C^{(k)}} |w|,$$

so gilt

$$\lim_{\lambda \rightarrow \infty} \frac{1}{n_1(k)} \sum_{\mu=n_k}^{n_1(k)} \log |w_\mu^{(k)}| \geq \log m_k,$$

mithin ist

$$\overline{\lim}_{s < \mathfrak{B}_k^{(1,2)}} |\mathcal{P}^{(k)}(z)| \geq m_k$$

und deshalb um so mehr

$$(22) \quad \overline{\lim}_{s < \mathfrak{B}} \left(\overline{\lim}_{n \rightarrow \infty} \left| \sqrt[n]{\prod_{\mu=1}^n \frac{z - a_\mu}{z - b_\mu}} \right| \right) \geq m_k.$$

⁷⁾ Wegen eines Beweises vgl. E. Lammel, Über Approximation regulärer Funktionen eines komplexen Argumentes durch rationale Funktionen. Math. Zeitschr. 46 (1940), S. 111–112.

Die Folge der Bereiche $\{\mathfrak{B}_k'\}$, ($k = 1, 2, \dots$) gehe bei der Abbildung $t = \frac{1}{\tau}$ in $\{\mathfrak{B}_k^{(v)}\}$, ($k = 1, 2, \dots$) über, die Folge ihrer Berandungskurven $\{C_k\}$, ($k = 1, 2, \dots$) in $\{C_k^{(v)}\}$, ($k = 1, 2, \dots$) und C in $C^{(v)}$.

Nun ist $\mathfrak{B}_{k+1}^{(v)} + C_{k+1}^{(v)}$ in $\mathfrak{B}_k^{(v)}$ enthalten und die Kurvenfolge $\{C_k^{(v)}\}$, ($k = 1, 2, \dots$) approximiert gleichmäßig die Kurve $C^{(v)}$. Ist $w_k(\tau)$ die zu $\tau_k(w)$ inverse Funktion, so konvergiert die Folge der Abbildungsfunktionen $\{w_k(\tau)\}$, ($k = 1, 2, \dots$) auf $\mathfrak{B}^{(v)} + C^{(v)}$ gleichmäßig gegen $w(\tau)$, wobei $w(\tau)$ den Bereich $\mathfrak{B}^{(v)}$, der das Innere von $C^{(v)}$ ist, so auf $|w| < 1$ konform abbildet, daß $w(0) = 0$ und $w'(0) > 0$ ist⁸⁾.

Hieraus folgt

$$\lim_{k \rightarrow \infty} m_k = 1,$$

und aus (22) und (20) ergibt sich mithin die Richtigkeit von (21).

Auf dieselbe Art, wie man (14) aus (13) erschließt, läßt sich aus (21)

$$(23) \quad \lim_{n \rightarrow \infty} \left| \sqrt[n]{\prod_{\mu=1}^n \frac{t - a_\mu}{t - b_\mu}} \right| = 1$$

herleiten, und zwar gilt (23) für jede Stelle t aus \mathfrak{B}' .

$\log|t - z(w, a)|$ ist eine auf $|w| \leq 1$ stetige Funktion von w , wenn t eine feste Stelle aus \mathfrak{B}' ist und $z(w, a)$ die früher bereits verwendete Abbildungsfunktion bedeutet. (23) sagt aus, daß die Bildfolgen $\{\alpha_v\}$, ($v = 1, 2, \dots$) und $\{\beta_v\}$, ($v = 1, 2, \dots$) die Eigenschaft haben, daß

$$(24) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\mu=1}^n \log \left| \frac{t - z(\alpha_\mu, a)}{t - z(\beta_\mu, a)} \right| = 0$$

gilt.

Da sich aber sowohl Real- als auch Imaginärteil von w^k für jedes positive ganzzahlige k auf $|w| \leq 1$ gleichmäßig durch Ausdrücke von der Form

$$\sum_{m=1}^l B_m \log |t_m - z(w, a)|, \quad ^9)$$

wobei die B_m Konstante und die Stellen t_m im Äußern von C gelegen sind, approximieren lassen, so folgt mit Hilfe von (24), daß die Bildfolgen $\{\alpha_\mu\}$, ($\mu = 1, 2, \dots$) und $\{\beta_\mu\}$, ($\mu = 1, 2, \dots$) den Bedingungen (3) genügen müssen. Diese sind also notwendig dafür, daß sich jede Funktion $f(z)$ der Klasse $H(b)$ in eine Reihe (2) entwickeln läßt.

⁸⁾ Wegen eines Beweises vgl. z. B. das in Fußnote ¹⁾ zitierte Werk, S. 32–35.

⁹⁾ Vgl. das in Fußnote ¹⁾ zitierte Werk, S. 170.

Untersuchungen zum Erneuerungsproblem.

Von

Hans Richter in Leipzig *).

Einleitung.

Das Erneuerungsproblem entstammt dem Fragenkreis der Versicherung, wo man bei der Betrachtung von Kassen vor allem sozialversicherungsähnlicher Natur darauf geführt wird¹⁾. Nachdem dann einmal seine mathematische Formulierung gefunden war, wurde ihm in der versicherungsmathematischen Literatur vor allem seit Chr. Moser eine größere Beachtung zuteil.

In der vorliegenden Arbeit findet man die genaue Problemstellung im § 1. Bezüglich der dort erklärten Absterbeordnung $P(t)$ sei noch darauf hingewiesen, daß man dieselbe diskontinuierlich als Treppenfunktion mit äquidistanten Sprüngen²⁾ oder wie hier als stetig betrachten kann. Weiter unterscheidet man je nachdem, ob $P(t)$ schon von einem endlichen Werte ω an verschwindet oder nicht, den Fall des endlichen und den des unendlichen „Höchstalters“ ω .

Den Ausgangspunkt der Untersuchung bildet die im § 2 unter allgemeineren als den üblich gemachten Voraussetzungen abgeleitete Integralgleichung (5), deren Lösung $\varphi(t)$ vor allem hinsichtlich ihres Verhaltens für große t untersucht werden soll. In der Literatur findet man nun die Vermutung, daß $\varphi(t)$ für große t gegen einen Grenzwert strebt. Dies trifft für den Fall des endlichen Höchstalters auch tatsächlich zu³⁾. Der Fall des unendlichen Höchstalters dagegen, für den bezüglich der Konvergenz von $\varphi(t)$ in der Literatur ein Gegenbeispiel von H. Hadwiger⁴⁾ vorhanden ist, soll hier geklärt werden; selbstverständlich wird dabei aber der Fall des endlichen Höchstalters nochmals mit erfaßt.

Da die Integralgleichung (5) vom Faltungstypus ist, wird die Lösung im § 4 durch die Methode der Fourierschen Transformierten angesetzt, mit dem Ziel, aus dem Verhalten der Transformierten von $\varphi(t)$ Rückschlüsse auf das Verhalten von $\varphi(t)$ zu ziehen⁵⁾. Die bei diesem Verfahren notwendige

*) D. 15. Eingereicht zur Erlangung des Grades eines Dr. phil. habil. in der Philosophischen Fakultät der Universität Leipzig.

¹⁾ Vgl. die im Literaturverzeichnis unter III angegebenen Arbeiten.

²⁾ Vgl. Literaturverzeichnis unter II.

³⁾ Richter (1). (Die Zitate beziehen sich auf das Literaturverzeichnis I.)

⁴⁾ Hadwiger (2).

⁵⁾ Sogenannte indirekte Abelsche Asymptotik; vgl. Doetsch, S. 224.

Verschiebung des Integrationsweges bei der Umkehrformel auf den Rand der Regularitätshalbebene wurde jedoch erst nach Zerlegung der komplexen Umkehrformel in ihre reellen Bestandteile durchgeführt; es zeigt sich nämlich, daß diese Verschiebung bei einem dieser Bestandteile im Gegensatz zur komplexen Formel auch noch bei den hier vorkommenden Singularitäten möglich ist (§ 5). Auf dem angegebenen Wege gelingt es, unter ziemlich allgemein gehaltenen Voraussetzungen zu zeigen, wann $\varphi(t)$ einem endlichen Grenzwert zustrebt und wann nicht (§ 6).

Im § 7 werden die Ergebnisse auf den Fall übertragen, daß die Anfangsaltergliederung und der vorgeschriebene Bestand der betrachteten Gesamtheit beliebig sind.

Das in der vorliegenden Untersuchung betrachtete wirkliche Geschehen wird durch Wahrscheinlichkeitsgesetze beherrscht. Der deterministische Ansatz durch eine Integralgleichung entspricht aber der typisch versicherungsmathematischen Betrachtungsweise, zunächst sogenannten „rechnungsmäßigen Ablauf“ anzunehmen, d. h. vorauszusetzen, daß beobachtete Häufigkeiten mit den ihnen zugrunde liegenden Wahrscheinlichkeiten zusammenfallen. Jedoch wird im § 3 und § 8 der Tatsache des Wahrscheinlichkeitscharakters aller Aussagen durch exakte Berechnung von Erwartungswert und Streuung Rechnung getragen, wobei sich zeigt, daß unter geringen Voraussetzungen der Quotient aus Streuung und Erwartungswert mit wachsendem t tatsächlich gegen Null läuft.

Die dem reinen Mathematiker im allgemeinen nicht geläufigen Grundbegriffe der Versicherungsmathematik werden im § 1 kurz angegeben. In mathematischer Hinsicht ist die Darstellung ausführlicher als üblich gehalten, damit die Arbeit auch für den angewandten Mathematiker, insbesondere Versicherungsmathematiker, lesbar bleibt, der die Theorie der Fourierschen Transformaten nicht beherrscht. Vielleicht kann so mit geholfen werden, daß dieses wichtige mathematische Hilfsmittel in weiteren Kreisen der angewandten Mathematik Allgemeingut wird.

§ 1.

Problemstellung.

Der Beobachtung mögen gewisse Elemente (etwa Personen) zugrunde liegen, wobei jedem Element eine bestimmte Eigenschaft x mit $0 \leq x < \infty$, das „Alter“ des Elementes, zukommt. Unterliegt ein Element des Alters x während der Zeitspanne t der Beobachtung, so gehe sein Alter in den Wert $x + t$ über. Der Einfachheit halber werden wir ein Element der Eigenschaft x als x -jährig bezeichnen.

Zu jedem Alter x und der Zeit t soll weiter eine Wahrscheinlichkeit $P(x, t)$ dafür bestehen, daß ein x -Jähriger nach der Zeit t noch nicht aus der Beobachtung ausgeschieden, „gestorben“, ist. $P(x, t)$ soll dabei der Bedingung

$$P(0, t) \cdot P(t, t_1) = P(0, t + t_1)$$

unterliegen, so daß

$$P(x, t) = \frac{P(x+t)}{P(x)}$$

bei $P(t) = P(0, t)$ wird.

$P(t)$ heißt die Absterbeordnung, die als eine nichtsteigende Funktion mit $P(0) = 1$ und $\lim_{t \rightarrow \infty} P(t) = 0$ vorausgesetzt sei. Üblicherweise fordert man noch, daß ein mittleres Alter $\bar{x} = \int_0^{\infty} P(t) dt$ existiert. Wegen des monoton gegen 0 gehenden Verhaltens von $P(t)$ ist für die Existenz von $\bar{x} \lim_{t \rightarrow \infty} t \cdot P(t) = 0$ notwendig, jedoch nicht hinreichend.

Gewöhnlich setzt man $P(t)$ noch als differenzierbar mit stückweise stetiger Ableitung voraus. Wir wollen diese Voraussetzung abschwächen zu der Forderung, daß es lediglich eine L -integrierbare nichtnegative Funktion $p(t)$ derart gebe, daß

$$(1) \quad P(t) = \int_t^{\infty} p(t) dt$$

wird, daß also $P(t)$ totalstetig ist. $p(t)$ wird die Ausscheideintensität genannt und soll in jedem endlichen Intervall beschränkt sein:

$$(2) \quad 0 \leq p(\tau) \leq M(t) \quad \text{für } 0 \leq \tau \leq t.$$

Die bekannte Sterbensintensität ist dann

$$\mu(t) = \frac{p(t)}{P(t)}.$$

Da für fast alle t — d. h. für alle bis auf eine Menge vom Lebesgueschen Maße Null — $P(t)$ auch im gewöhnlichen Sinne die Ableitung $p(t)$ besitzt, hat für fast alle x $\mu(x)$ die Bedeutung, daß $\mu(x) dt$ die Wahrscheinlichkeit angibt, daß ein x -Jähriger im Laufe der folgenden Zeit dt ausscheidet. Für eine Restmenge vom Maße 0 ist $\mu(x)$ beliebig wählbar; auf Grund der über $p(t)$ gemachten Voraussetzungen können wir jedenfalls annehmen, daß $\mu(x)$ für kleine x beschränkt bleibt. Falls $P(t) = 0$ nicht schon von einem endlichen t -Wert ab erfüllt ist, kann $\mu(x)$ natürlich als in jedem endlichen Intervall beschränkt angesehen werden.

Mit $P(t)$ ist in einem Intervall $0 \leq x \leq A$ bei $P(A) > 0$ auch $\ln P(t)$ totalstetig und daher

$$\ln P(t) = \int_0^t \frac{P'(t)}{P(t)} dt$$

oder

$$P(t) = e^{-\int_0^t \mu(t) dt}.$$

Wir bemerken, daß die oben gemachte Einschränkung, daß $\int_0^\infty P(t) dt$ endlich sein soll, in allen praktischen Fällen erfüllt sein wird. Bei endlichem Höchstalter liegt kein Zweifel vor. Bei unendlichem Höchstalter wird man insbesondere solche Fälle zu betrachten haben, wo die Sterbensintensität natürlicherweise mit wachsendem Alter steigt. Dieser Fall wird aber erfaßt durch

Satz 1. Ist $\lim_{x \rightarrow \infty} x \cdot \mu(x) > l$, so ist $\int_0^\infty x^n p(x) dx$ für alle $n \leq l$ konvergent.

Beweis. Nach Voraussetzung gibt es ein A so, daß für $x > A$

$$\mu(x) \geq \frac{l'}{x}$$

gilt, wo $l' > l$ ist. Für $x > A$ ist somit

$$P(x) = P(A) e^{-\int_A^x \mu(x) dx} \leq P(A) e^{-\int_A^x \frac{l'}{x} dx} = P(A) \frac{A^{l'}}{x^{l'}}$$

und hieraus

$$\int_0^\infty P(x) x^{n-1} dx \leq \int_0^A P(x) x^{n-1} dx + P(A) \cdot A^{l'} \cdot \int_A^\infty x^{n-l'-1} dx,$$

was für $n \leq l$ wegen $l < l'$ sicher konvergiert. Mit $\int_0^\infty P(x) \cdot x^{n-1} dx$ konvergiert aber wegen des monotonen Verhaltens von $P(x)$ auch $\int_0^\infty p(x) \cdot x^n dx$.

Wenn, wie in praktischen Fällen, $\lim_{x \rightarrow \infty} \mu(x) > 0$ ist, ist also $\int_0^\infty p(x) \cdot x^n dx$ für alle n konvergent.

Vorgegeben sei nunmehr eine als aus kontinuierlich vielen Elementen bestehend gedachte Gesamtheit, wobei jedes Element einer Ausscheidung der besprochenen Art bei rechnungsmäßigem⁶⁾ Verlauf unterliegen soll.

⁶⁾ Vgl. die Einleitung.

Zur Zeit $t = 0$ möge die Gesamtheit den Umfang $H(0) = 1$ besitzen und aus lauter Elementen des Alters Null bestehen⁷⁾. Außerdem soll die Gesamtheit einen laufenden Zugang an 0-Jährigen erleben, durch den der Gesamtheit bis zum Zeitpunkt t insgesamt $\Phi(t)$ neue Mitglieder zugeführt werden. $\Phi(t)$ ist naturgemäß eine positive, nichtfallende Funktion mit $\Phi(0) = 0$. Der Bestand der Gesamtheit zur Zeit t beträgt dann

$$(3) \quad H(t) = P(t) + \int_0^t P(t-k) d\Phi(k).$$

Das Erneuerungsproblem besteht nun in der Aufgabe, die Zugangsfunktion $\Phi(t)$ derart zu bestimmen, daß der Bestand der Gesamtheit konstant bleibt; insbesondere ist zu untersuchen, ob für große t eine Stabilisierung eintritt derart, daß $\frac{1}{t} \Phi(t)$ gegen eine Konstante konvergiert. Ein solches Vorkommnis werden wir „Stabilisierung im Mittel“ nennen. Außerdem — und das ist die in der Literatur gewöhnlich gestellte Aufgabe — werden wir zu untersuchen haben, ob vielleicht $\Phi(t)$ sogar eine Ableitung besitzt, die für große t gegen eine Konstante strebt: „Eigentliche Stabilisierung“. Im Fall der Existenz einer Ableitung $\varphi(t)$ der Zugangsfunktion $\Phi(t)$ nennen wir $\varphi(t)$ die „Erneuerungsintensität“ oder die „Erneuerungsfunktion“, die zur Absterbeordnung $P(t)$ und zur Ausscheideintensität $p(t)$ gehört. $\varphi(t)$ wird selbstverständlich durch das Problem nur bis auf eine Nullfunktion bestimmt.

Gehen wir jetzt wieder von (3) aus und machen über $\Phi(t)$ die Annahme, daß es in jedem endlichen Intervall beschränkt bleibt, dann zieht dies natürlich die gleiche Eigenschaft für $H(t)$ nach sich. Dieser Zusammenhang gilt aber auch in der umgekehrten Richtung, so daß man bei der Lösung von (3) bei beliebigen, aber in jedem endlichen Intervall beschränkten $H(t)$ von vornherein nur solche $\Phi(t)$ anzusetzen braucht, die für endliche t -Intervalle beschränkt sind. Der Beweis für diese Tatsache ergibt sich leicht folgendermaßen:

Sei $\delta \geq 0$ so gewählt, daß noch $P(\delta) > 0$ ist. Dann ist nach (3)

$$H(t) = P(t) + \int_0^{t-\delta} P(t-k) d\Phi(k) + \int_{t-\delta}^t P(t-k) d\Phi(k),$$

also

$$\int_{t-\delta}^t P(t-k) d\Phi(k) = H(t) - P(t) - \int_0^{t-\delta} P(t-k) d\Phi(k) < H(t).$$

⁷⁾ Wenn in einem praktischen Fall, z. B. einer Sterbekasse, die Mitgliedschaft ein Mindestalter voraussetzt, so ist unter x das um den Mindestbetrag verminderte Lebensalter zu verstehen.

Andererseits ist

$$\int_{t-\delta}^t P(t-k) d\Phi(k) > \int_{t-\delta}^t P(\delta) d\Phi(k) = P(\delta) \cdot [\Phi(t) - \Phi(t-\delta)]$$

und damit

$$\Phi(t) - \Phi(t-\delta) < \frac{H(t)}{P(\delta)}$$

für alle t , woraus die Behauptung unmittelbar abzulesen ist.

§ 2.

Die Lösbarkeit der Aufgabe.

(3) ist als Volterrasche Integralgleichung erster Art und wegen des in ihr vorkommenden Stieltjesschen Integrals für die Behandlung etwas ungeeignet. Zum Zwecke der späteren Umformung in eine gewöhnliche Gleichung zweiter Art zeigen wir daher zunächst

Satz 2. Ist $H(t)$ totalstetig in $0 \leq t \leq T$ und $\Phi(t)$ in diesem Intervall Lösung von (3), so ist auch $\Phi(t)$ dort totalstetig.

Beweis. Es sei δ eine kleine positive Zahl mit $P(\delta) > \frac{1}{2}$, \mathfrak{M} eine Menge von endlich vielen in $0 \leq t \leq T$ gelegenen Intervallen $a_i \leq t \leq b_i$ mit $\sum (b_i - a_i) < \delta$. Ist nun $\Phi(t)$ Lösung von (3), so gilt

$$\begin{aligned} H(b_i) - H(a_i) &= [P(b_i) - P(a_i)] + \int_0^{a_i} [P(b_i - k) - P(a_i - k)] d\Phi(k) + \\ &\quad + \int_{a_i}^{b_i} P(b_i - k) d\Phi(k) \end{aligned}$$

oder

$$\begin{aligned} \int_{a_i}^{b_i} P(b_i - k) d\Phi(k) &= [H(b_i) - H(a_i)] + [P(a_i) - P(b_i)] + \\ &\quad + \int_0^{a_i} [P(a_i - k) - P(b_i - k)] d\Phi(k). \end{aligned}$$

Nach Voraussetzung ist weiter $P(\delta) > \frac{1}{2}$, also

$$\frac{1}{2} [\Phi(b_i) - \Phi(a_i)] < \int_{a_i}^{b_i} P(\delta) d\Phi(k) \leq \int_{a_i}^{b_i} P(b_i - k) d\Phi(k),$$

und endlich wegen der aus (2) folgenden Relation

$$0 \leq P(a_i) - P(b_i) \leq M(T) (b_i - a_i)$$

haben wir

$$\begin{aligned} \frac{1}{2} \sum |\Phi(b_i) - \Phi(a_i)| &< \sum |H(b_i) - H(a_i)| + M(T) \cdot \sum (b_i - a_i) + \\ &\quad + M(T) \Phi(T) \sum (b_i - a_i) \\ &\leq \sum |H(b_i) - H(a_i)| + \delta M(T) [1 + \Phi(T)]. \end{aligned}$$

Also

$$\lim_{\delta \rightarrow 0} \sum |\Phi(b_i) - \Phi(a_i)| = 0.$$

Wenn aber eine Funktion für jede endliche Intervallmenge totalstetig ist, so ist sie bekanntlich⁹⁾ überhaupt totalstetig. Da andererseits Totalstetigkeit weiter nichts als Darstellbarkeit in Form eines unbestimmten Integrals bedeutet⁹⁾, können wir das gewonnene Ergebnis auch so formulieren:

Satz 2a. Ist $H(t) = 1 + \int_0^t h(t) dt$ und $\Phi(t)$ Lösung von (3), so gibt es ein $\varphi(t)$ mit $\Phi(t) = \int_0^t \varphi(t) dt$.

In genau der gleichen Weise wie Satz 2 zeigt man

Satz 2b. Ist $H(t)$ in $0 \leq t \leq T$ von beschränktem Differenzenquotienten und $\Phi(t)$ Lösung von (3), so ist auch $\Phi(t)$ dort von beschränktem Differenzenquotienten.

In den damit erfaßten Fällen wie etwa dem des konstanten Gesamtbestandes ist also von vornherein klar, daß die Zugangsintensität existiert und in jedem endlichen Intervall als beschränkt angenommen werden kann, falls (3) überhaupt lösbar ist, was wir damit natürlich noch nicht gezeigt haben.

Ist nun in der Tat $\Phi(t)$ totalstetig, also $\Phi(t) = \int_0^t \varphi(t) dt$, dann gilt¹⁰⁾ für das in (3) stehende Integral

$$\int_0^t P(t-k) d\Phi(k) = \int_0^t P(t-k) \varphi(k) dk.$$

Weiter ist aber mit Hilfe einer nach Fubini¹¹⁾ gerechtfertigten Vertauschung der Integrationsreihenfolge

$$\begin{aligned} \int_0^t P(t-k) \varphi(k) dk &= \int_0^t \varphi(k) \left[1 - \int_0^{t-k} p(l) dl \right] dk \\ &= \int_0^t \varphi(k) dk - \int_0^t \varphi(k) \left[\int_k^t p(l-k) dl \right] dk \\ &= \int_0^t \varphi(k) dk - \int_0^t \left[\int_0^l p(l-k) \varphi(k) dk \right] dl. \end{aligned}$$

⁹⁾ Vgl. Carathéodory, § 457, Satz 2.

⁹⁾ Vgl. Carathéodory, § 440, Satz 6.

¹⁰⁾ Vgl. Saks, S. 97, Theorem 18.

¹¹⁾ Vgl. Carathéodory, § 548–554.

Bis auf eine Ausnahmemenge vom Maße Null folgt damit aus (3)

$$h(t) = -p(t) + \varphi(t) - \int_0^t p(t-k) \varphi(k) dk$$

oder

$$(4) \quad \varphi(t) = h(t) + p(t) + \int_0^t p(t-k) \varphi(k) dk$$

für fast alle t . Da aber andererseits $\varphi(t)$ durch das Problem nur bis auf Äquivalenz bestimmt ist, können wir $\varphi(t)$ unbedenklich so normieren, daß (4) für alle t gilt.

Wir fassen zusammen zu

Satz 3. Ist $H(t) = 1 + \int_0^t h(t) dt$ und $\Phi(t)$ Lösung von (3), dann ist $\Phi(t) = \int_0^t \varphi(t) dt$, wo $\varphi(t)$ der Gleichung

$$\varphi(t) = h(t) + p(t) + \int_0^t p(t-k) \varphi(k) dk$$

genügt.

Für das eigentliche Erneuerungsproblem genügt $\varphi(t)$ insbesondere der Integralgleichung

$$(5) \quad \varphi(t) = p(t) + \int_0^t p(t-k) \varphi(k) dk.$$

Für ein Intervall $0 \leq t \leq T$ ist $p(t) \leq M(T)$ und daher nach (5)

$$|\varphi(t)| \leq M(T) \cdot \left[1 + \int_0^t |\varphi(k)| dk \right],$$

woraus wir die Folgerung ziehen¹²⁾

$$(6) \quad |\varphi(t)| \leq M(T) \cdot e^{t \cdot M(T)} \text{ in } 0 \leq t \leq T.$$

Bilden wir jetzt eine Folge von Funktionen durch die Festsetzung

$$p_1(t) = p(t)$$

$$p_n(t) = \int_0^t p_{n-1}(k) p(t-k) dk,$$

dann sind die nirgends negativen $p_n(t-k)$ ersichtlich die iterierten Kerne der Integralgleichungen (4) und (5), die der aus $p(t) \leq M(T)$ in $0 \leq t \leq T$ durch vollständige Induktion leicht beweisbaren Abschätzung

$$(7) \quad p_n(t) \leq M^n(T) \cdot \frac{t^{n-1}}{(n-1)!}$$

genügen.

¹²⁾ Vgl. Richter (1), S. 25.

Um den lösenden Kern unserer Integralgleichungen zu erhalten, bilden wir jetzt die in jedem endlichen Intervall $0 \leq t \leq T$ wegen (7) sicher gleichmäßig konvergente Summe

$$\psi(t) = \sum_1^{\infty} p_n(t).$$

Zur Berechnung von $\int_0^t p(t-k) \psi(k) dk$ können wir wegen der Gleichmäßigkeit der Konvergenz die Integration gliedweise ausführen¹³⁾ und erhalten

$$\int_0^t p(t-k) \psi(k) dk = \sum_1^{\infty} \int_0^t p(t-k) p_n(k) dk = \sum_2^{\infty} p_n(t) = \psi(t) - p(t),$$

was $\psi(t)$ als Lösung von (5) erweist. Dann ist aber

$$\begin{aligned} & [\psi(t+\delta) - p(t+\delta)] - [\psi(t) - p(t)] \\ & \leq \int_0^t |p(t+\delta-k) - p(t-k)| \psi(k) dk + \int_t^{t+\delta} p(t+\delta-k) \psi(k) dk, \end{aligned}$$

woraus wir wegen der Beschränktheit von $\psi(t)$ in jedem endlichen Intervall und dem bekannten Lemma von Lebesgue¹⁴⁾

$$\lim_{\delta \rightarrow 0} \int_0^t |p(k+\delta) - p(k)| dk = 0$$

die Stetigkeit von $\psi(t) - p(t)$ schließen können. Es ist also

$$(8) \quad \psi(t) = p(t) + r(t)$$

mit stetigem $r(t)$.

Bilden wir weiter die Funktion

$$(9) \quad \chi(t) = \psi(t) + h(t) + \int_0^t h(k) \psi(t-k) dk,$$

so können wir wegen der Beschränktheit der Partialsummen $\sum_1^m p_n(t)$ dafür schreiben

$$\chi(t) = \psi(t) + h(t) + \sum_1^{\infty} \int_0^t h(k) p_n(t-k) dk$$

¹³⁾ Carathéodory, § 402, Satz 16.

¹⁴⁾ Lebesgue (1), S. 15.

und erhalten hieraus¹⁵⁾

$$\begin{aligned} \int_0^t p(t-k) \chi(k) dk &= \int_0^t p(t-k) \psi(k) dk + \int_0^t p(t-k) h(k) dk + \\ &\quad + \sum_1^\infty \int_{k=0}^t \int_{l=0}^k p(t-k) h(l) p_n(k-l) dl dk \\ &= \psi(t) - p(t) + \int_0^t p(t-k) h(k) dk + \sum_1^\infty \int_{l=0}^t h(l) \int_{k=l}^t p_n(k-l) p(t-k) dk dl \\ &= \psi(t) - p(t) + \int_0^t p(t-k) h(k) dk + \sum_1^\infty \int_0^t h(l) p_{n+1}(t-l) dl \end{aligned}$$

oder gemäß (9)

$$\int_0^t p(t-k) \chi(k) dk = \chi(t) - p(t) - h(t),$$

was $\chi(t)$ als Lösung von (4) erweist.

Wir fassen zusammen zu

Satz 4. Ist $P(t) = 1 - \int_0^t p(t) dt$ mit $p(t) \geq 0$ und $\int_0^\infty p(t) dt = 1$, so ist das Erneuerungsproblem lösbar. Lösung ist eine Funktion $\psi(t)$ mit den Eigenschaften: $\psi(t) \geq 0$, $\psi(t) - p(t)$ ist stetig, $\psi(t) \leq M(t) \cdot e^{N(t)}$, $\psi(t-k)$ ist der lösende Kern zu der zum erweiterten Erneuerungsproblem gehörigen Integralgleichung (4).

Die Eindeutigkeit der Lösbarkeit von (5) könnte an dieser Stelle ebenfalls bewiesen werden¹⁶⁾; sie wird sich jedoch später allein ergeben.

Wir sind bereits jetzt in der Lage, eine notwendige Bedingung dafür anzugeben, daß eigentliche Stabilisierung eintritt. Durch diese Bedingung wird insbesondere das in der Einleitung erwähnte Gegenbeispiel von Hadwiger¹⁷⁾ ausgeschlossen. Es gilt nämlich¹⁸⁾

Satz 5. Notwendig dafür, daß $\varphi(t)$ gegen einen Grenzwert c strebt, ist die Existenz einer äquivalenten Funktion $\hat{p}(t)$ zu $p(t)$ mit $\lim_{t \rightarrow \infty} \hat{p}(t) = 0$.

¹⁵⁾ Vgl. dabei Anm. 13.

¹⁶⁾ Da die Differenz zweier Lösungen stetig sein muß, kann man nämlich den üblichen Eindeutigkeitsbeweis führen.

¹⁷⁾ Hadwiger (2).

¹⁸⁾ Vgl. Richter (2).

Für den Beweis benötigen wir noch den in den weiteren Untersuchungen oft zu verwendenden Hilfssatz

$$(10) \left\{ \begin{array}{l} \text{Sind } f(x) \text{ und } g(x) \text{ integrierbar und gilt entweder} \\ a) \int_0^{\infty} |f(x)| dx \text{ ist konvergent und } g(x) \text{ beschränkt mit } \lim_{x \rightarrow \infty} g(x) = 0, \text{ oder} \\ b) \int_0^{\infty} f^2(x) dx \text{ und } \int_0^{\infty} g^2(x) dx \text{ sind beide konvergent; dann ist} \\ \lim_{x \rightarrow \infty} \int_0^x f(x-k) g(k) dk = 0. \end{array} \right.$$

Zum Beweis zerlegt man $\int_0^x f(x-k) g(k) dk$ in die Summanden $\int_0^{\frac{x}{2}} f(x-k) g(k) dk$ und $\int_{\frac{x}{2}}^x f(x-k) g(k) dk$, die sich auf Grund der gemachten

Voraussetzungen — im Falle b) mit Hilfe der Schwarzischen Ungleichheit — unmittelbar abschätzen lassen.

Wir kommen nun zum

Beweis von Satz 5: Nach Satz 3 ist $\varphi(t) \sim \widehat{\varphi}(t)$, wo $\widehat{\varphi}(t)$ der Gleichung (5) genügt. Wegen der für $\widehat{\varphi}(t)$ gültigen Abschätzung (6) und dem Grenzwertverhalten von $\varphi(t)$ kann $\varphi(t)$ als beschränkt angesehen werden. Sei nun in der Tat $\varphi(t) = c + \gamma(t)$ mit $\lim_{t \rightarrow \infty} \gamma(t) = 0$, so ist auch $\gamma(t)$ beschränkt; es wird daher nach (10)

$$\lim_{t \rightarrow \infty} \int_0^t p(t-k) \gamma(k) dk = 0.$$

Also ist

$$\begin{aligned} \lim_{t \rightarrow \infty} \int_0^t p(t-k) \widehat{\varphi}(k) dk &= \lim_{t \rightarrow \infty} \int_0^t p(t-k) \varphi(k) dk \\ &= \lim_{t \rightarrow \infty} \int_0^t c \cdot p(k) dk + \lim_{t \rightarrow \infty} \int_0^t p(t-k) \gamma(k) dk = c + 0 = c. \end{aligned}$$

Nach (5) ist nun

$$p(t) = \widehat{\varphi}(t) - \int_0^t p(t-k) \widehat{\varphi}(k) dk = \varphi(t) - \int_0^t p(t-k) \widehat{\varphi}(k) dk + [\widehat{\varphi}(t) - \varphi(t)],$$

also

$$p(t) \sim \widehat{p}(t)$$

mit

$$\widehat{p}(t) = \varphi(t) - \int_0^t p(t-k) \widehat{\varphi}(k) dk.$$

Dabei ist wie behauptet

$$\lim_{t \rightarrow \infty} \hat{p}(t) = \lim_{t \rightarrow \infty} \varphi(t) - \lim_{t \rightarrow \infty} \int_0^t p(t-k) \hat{\varphi}(k) dk = c - c = 0.$$

§ 3.

Die wahrscheinlichkeitstheoretische Betrachtungsweise.

Bei der von uns angeführten Problemstellung im § 1 sind wir davon ausgegangen, daß das Absterben der Gesamtheit genau nach der vorgelegten Absterbeordnung erfolgt. Diese Darstellung erscheint solange haltbar, als es sich — wie oben auch vorausgesetzt wurde — um eine aus kontinuierlich vielen Elementen bestehende Gesamtheit handelt. Nichtsdestoweniger verdient gerade von einem wirklichkeitsnäheren Standpunkt aus die Frage Beachtung, inwieweit die Ergebnisse der im § 1 gestellten Aufgabe auch auf eine Gesamtheit angewendet werden können, die aus endlich vielen sich gegenseitig nicht beeinflussenden diskreten Elementen zusammengesetzt ist. Es wird dabei selbstverständlich genügen, eine Gesamtheit zu betrachten, die nur aus einem einzigen Mitgliede besteht, für das beim Ableben 0-jähriger Ersatz geschafft wird.

Da die Erneuerungsintensität dementsprechend nur noch die Bedeutung einer Wahrscheinlichkeitsdichte haben kann, werden wir zur Vermeidung unnötig komplizierender Details in diesem den Gang der übrigen Untersuchung nicht beeinflussenden Paragraphen voraussetzen, daß $\mu(t)$ und damit auch $p(t)$ für alle t die Bedeutung von Wahrscheinlichkeitsdichten haben, um eine bequeme infinitesimale Betrachtungsweise zu ermöglichen. Dann ist aber völlig klar, daß $p_n(t) dt$ die Wahrscheinlichkeit dafür ist, daß das Element zur Zeit t bis $t + dt$ nach $(n-1)$ -maligem vorherigem Ersatz zum n -ten Male ausscheidet.

$$\psi(t) = \sum_1^{\infty} p_n(t)$$

ist also in der Tat die Ausscheidewahrscheinlichkeitsdichte zur Zeit t ¹⁹⁾.

Ziel der folgenden Betrachtung ist es nun, zunächst zu zeigen, daß auch den Integralgleichungen (3) bei $H(t) \equiv 1$ und (5) eine wahrscheinlichkeitstheoretische Bedeutung zukommt, und schließlich die Frage zu klären, welche Streuung die Anzahl der wirklich eingetretenen Erneuerungen des Elementes um ihren Erwartungswert zur Zeit t hat, was zu einem neuen Stabilitätsbegriff führt. Die hier behandelte Streuung ist wohl zu unterscheiden von der Streuung, auf die die Frage nach der Konkretisierung des Erwartungswertes

¹⁹⁾ Vgl. Kohler.

bei Verwendung einer genügend großen Anzahl von Elementen führt¹⁰⁾, was wahrscheinlichkeitstheoretisch nichts Neues bieten kann.

Es sei nun t ein fest gewählter Zeitpunkt, x ein Alter mit $x < t$, $q_n(x, t) dx$ die Wahrscheinlichkeit dafür, daß das Element bis zur Zeit t genau n -mal ausgeschieden ist und nunmehr das Alter x bis $x + dx$ besitzt. Es ist dann

$$q_0(x, t) = 0,$$

$$q_1(x, t) = p(t - x) P(x),$$

allgemein

$$q_n(x, t) = \int_{0 < \xi_1 < \dots < \xi_{n-1} < t-x} \dots \int p(\xi_1) p(\xi_2 - \xi_1) \dots p(\xi_{n-1} - \xi_{n-2}) \times \\ \times p(t - x - \xi_{n-1}) P(x) d\xi_1 \dots d\xi_{n-1}$$

oder

$$q_n(x, t) = p_n(t - x) P(x) \text{ für } n = 1, 2, \dots, q_0(x, t) = 0.$$

Die Wahrscheinlichkeit $w(x, t) dx$, das Alter x bis $x + dx$ zu besitzen, beträgt dann

$$w(x, t) = \sum_0^{\infty} q_n(x, t) = P(x) \sum_1^{\infty} p_n(t - x).$$

Mit $\psi(t) = \sum_1^{\infty} p_n(t)$ ist also

$$w(x, t) = \psi(t - x) P(x).$$

Die Wahrscheinlichkeit, genau das Alter t zu besitzen, ist selbstverständlich $w(t, t) = P(t)$. Da das Element zur Zeit t irgendein Alter haben muß, gilt

$$w(t, t) + \int_0^t w(x, t) dx = 1$$

oder

$$P(t) + \int_0^t \psi(t - x) P(x) dx = 1,$$

was weiter nichts als Gleichung (3) darstellt, die ja in der Tat von $\psi(t)$ erfüllt wird.

Für die Wahrscheinlichkeit $\psi(t) dt$ dafür, daß das Element in der Zeit t bis $t + dt$ ausscheidet, gilt offenbar weiter

$$\psi(t) = w(t, t) P(t) + \int_0^t w(x, t) P(x) dx \\ = P(t) + \int_0^t \psi(t - x) P(x) dx,$$

womit (5) wiedergefunden ist.

Suchen wir weiter die Wahrscheinlichkeit w_n dafür, daß das Element bis zum Zeitpunkt t genau n -mal erneuert wurde, so erhalten wir

$$\begin{cases} \text{bei } n = 0 & w_0 = P(t) \\ \text{bei } n \geq 1 & w_n = \int_0^t q_n(x, t) dx = \int_0^t p_n(t-x) P(x) dx. \end{cases}$$

Wir bezeichnen jetzt in Analogie zu § 1 mit $\Phi(t)$ die — notwendig ganzzahlige — Anzahl der Erneuerungen, die das Element bis zum Zeitpunkt t erlebt hat; w_n ist also die Wahrscheinlichkeit dafür, daß diese „Erneuerungszahl“ zur Zeit t den Wert n hat.

Da $\psi(t) dt$ der Erwartungswert für den Zuwachs von $\Phi(t)$ in der Zeit von t bis $t + dt$ ist, muß der Erwartungswert $\overline{\Phi(t)}$ von $\Phi(t)$ natürlich durch

$$(11) \quad \overline{\Phi(t)} = \int_0^t \psi(t) dt = \Psi(t)$$

gegeben sein. Es ist jedoch von Interesse, daß diese Gleichung bei Kenntnis der w_n nunmehr ohne weitere infinitesimale Betrachtung gefunden werden kann. Aus den w_n finden wir nämlich sofort

$$(12) \quad \overline{\Phi(t)} = \sum_0^\infty n w_n = \sum_1^\infty n \int_0^t p_n(t-x) P(x) dx,$$

und es kommt daher nur noch darauf an, (12) in (11) umzuformen.

Zuvor bemerken wir jedoch noch, daß bei der wahrscheinlichkeitstheoretischen Betrachtungsweise ein weiterer Stabilitätsbegriff für den Neuzugang aufgestellt werden kann. Wenn nämlich auch $\frac{1}{t} \overline{\Phi(t)}$ gegen eine Konstante konvergiert, so könnte doch vielleicht die Streuung $s(t)$ von $\Phi(t)$ ebenfalls von der Größenordnung t sein, so daß eine wirkliche Stabilisierung nicht zustande kommt. Wir wollen daher von „wahrscheinlichkeitstheoretischer Stabilisierung“ sprechen, wenn nicht nur $\frac{1}{t} \overline{\Phi(t)}$ gegen eine Konstante geht, sondern auch

$$\frac{s^2(t)}{\overline{\Phi(t)}^2} = \frac{[\overline{\Phi(t) - \Phi(t)}]^2}{\overline{\Phi(t)}^2} = \frac{\overline{\Phi^2(t)} - \overline{\Phi(t)}^2}{\overline{\Phi(t)}^2} \rightarrow 0$$

läuft. Dabei ist analog zu (12)

$$(13) \quad \overline{\Phi^2(t)} = \sum_1^\infty n^2 \int_0^t p_n(t-x) P(x) dx$$

zu setzen.

Um eine bequeme Berechnung der Ausdrücke (12) und (13) zu ermöglichen, gehen wir von den w_n zu der zugehörigen erzeugenden Funktion über. Bilden wir zu diesem Zwecke die für alle x und τ absolut konvergente Reihe

$$(14) \quad F(x, \tau) = \sum_1^{\infty} \tau^n p_n(x),$$

dann ist

$$\int_0^t F(t-x, \tau) P(x) dx = \sum_1^{\infty} \tau^n \int_0^t p_n(t-x) P(x) dx = \sum_1^{\infty} \tau^n w_n$$

und damit

$$(15) \quad \sum_0^{\infty} \tau^n w_n = P(t) + \int_0^t F(t-x, \tau) P(x) dx = e_t(\tau).$$

Wir haben also in $e_t(\tau)$ die erzeugende Funktion zu den Wahrscheinlichkeiten w_n gewonnen.

$\Phi(t)$ und $\Phi^2(t)$ drücken sich dann in bekannter Weise durch $e_t(\tau)$ folgendermaßen aus:

$$(16) \quad \begin{cases} \overline{\Phi(t)} = \sum_1^{\infty} n w_n = \left[\frac{\partial}{\partial \tau} e_t(\tau) \right]_{\tau=1} = e'_t(1), \\ \overline{\Phi^2(t)} = \sum_1^{\infty} n^2 w_n = \left[\frac{\partial^2}{\partial \tau^2} e_t(\tau) \right]_{\tau=1} + \left[\frac{\partial}{\partial \tau} e_t(\tau) \right]_{\tau=1} = e''_t(1) + e'_t(1). \end{cases}$$

Nun erkennen wir aber an (14) sofort, daß $F(t-x, \tau)$ der lösende Kern zum Kern $\tau p(t-x)$ ist. Die Gleichung (15) zeigt daher, daß $e_t(\tau)$ als Funktion von t der folgenden Integralgleichung Genüge leistet:

$$(17) \quad e_t(\tau) = P(t) + \tau \int_0^t e_k(\tau) p(t-k) dk.$$

Durch Differenzieren nach τ erhalten wir hieraus

$$(18) \quad e'_t(\tau) = \tau \int_0^t e'_k(\tau) p(t-k) dk + \int_0^t e_k(\tau) p(t-k) dk,$$

$$(19) \quad e''_t(\tau) = \tau \int_0^t e''_k(\tau) p(t-k) dk + 2 \int_0^t e'_k(\tau) p(t-k) dk.$$

Setzen wir in (17), (18) und (19) $\tau = 1$, so erhalten wir

$$e_t(1) = P(t) + \int_0^t e_k(1) p(t-k) dk,$$

$$e'_t(1) = \int_0^t e'_k(1) p(t-k) dk + \int_0^t e_k(1) p(t-k) dk,$$

$$e''_t(1) = \int_0^t e''_k(1) p(t-k) dk + 2 \int_0^t e'_k(1) p(t-k) dk,$$

oder damit gleichbedeutend unter Beachtung von (5)

$$(20) \quad e_i(1) \equiv 1,$$

$$(21) \quad e_i'(1) = \int_0^t e_i'(1) p(t-k) dk + 1 - P(t),$$

$$(22) \quad e_i''(1) = \int_0^t e_i''(1) p(t-k) dk + 2[e_i'(1) - 1 + P(t)].$$

Da $\psi(t-k)$ lösender Kern zum Kern $p(t-k)$ ist, lassen sich (21) und (22) unmittelbar lösen und ergeben unter Berücksichtigung von (3) bei $H(t) \equiv 1$:

$$\begin{aligned} e_i'(1) &= \int_0^t [1 - P(t)] \psi(t-k) dk + 1 - P(t) = \int_0^t \psi(k) dk = \Psi(t), \\ \frac{1}{2} e_i''(1) &= \Psi(t) - 1 + P(t) + \int_0^t [\Psi(k) - 1 + P(k)] \psi(t-k) dk \\ &= \Psi(t) - 1 + P(t) + \int_0^t \Psi(k) \psi(t-k) dk - \Psi(t) + 1 - P(t) \\ &= \int_0^t \Psi(k) \psi(t-k) dk. \end{aligned}$$

Wir erhalten somit aus (16)

$$\overline{\Phi}(t) = \Psi(t),$$

wie wir dies schon erwartet hatten, und

$$\overline{\Phi^2}(t) = 2 \int_0^t \Psi(k) \psi(t-k) dk + \Psi(t).$$

Das Quadrat der Streuung ist also

$$(23) \quad s^2(t) = 2 \int_0^t \Psi(k) \psi(t-k) dk + \Psi(t) - \Psi^2(t).$$

Auch für die Beurteilung der wahrscheinlichkeitstheoretischen Stabilisierung kommt es also darauf an, Kenntnis über das Verhalten von $\psi(t)$ bei großen t zu erlangen.

§ 4.

Die Lösung durch Fouriersche Integrale.

Da die Integralgleichung (5) vom sogenannten Faltungstypus²⁰⁾ ist, kann sie mit Hilfe der Fourierschen Transformaten behandelt werden²¹⁾.

²⁰⁾ Näheres hierüber siehe Doetsch, S. 279ff.

²¹⁾ Vgl. auch Kostitzin.

Führen wir nämlich zunächst einmal ohne Rücksicht auf Konvergenzfragen die Fourierschen Transformaten von $p(t)$ und $\varphi(t)$

$$(24) \quad \left\{ \begin{array}{l} f(z) = \int_0^{\infty} \varphi(t) e^{-izt} dt \\ \text{und} \\ b(z) = \int_0^{\infty} p(t) e^{-izt} dt \end{array} \right.$$

ein, so wird aus (5) auf Grund des Faltungssatzes²²⁾

$$f(z) = b(z) + f(z) b(z)$$

und damit

$$(25) \quad f(z) = \frac{b(z)}{1 - b(z)},$$

woraus sich unter passenden Voraussetzungen $\varphi(t)$ gemäß

$$(26) \quad \varphi(t) = \frac{1}{2\pi} \int_{-ia-\infty}^{-ia+\infty} f(z) e^{izt} dz$$

bei geeignetem a zurückberechnet.

Für eine exakte Betrachtungsweise ergeben sich hier jedoch bekanntlich die folgenden Teilaufgaben:

- a) Für welche z ist $b(z)$ definiert und regulär?
- b) Hat jede Lösung $\varphi(t)$ von (5) und insbesondere $\psi(t)$ überhaupt eine Transformatierte und stimmt diese mit $\frac{b(z)}{1-b(z)}$ überein?
- c) Läßt sich tatsächlich $\varphi(t)$ nach (26) zurückgewinnen?

a) Betrachtung von $b(z)$.

Da $p(t)$ absolut integrierbar ist, ist auf jeden Fall $b(z)$ in der unteren Halbebene regulär und einschließlich der reellen Achse stetig. Weiter brauchen wir an Kenntnissen über $b(z)$ nur noch den

Satz 6: $b(z) = 1$ hat in der unteren Halbebene einschließlich reeller Achse nur die Lösung $z = 0$.

Beweis. Bei $z = \xi - ia$ folgt zunächst aus (24)

$$(27) \quad |b(z)| \leq \int_0^{\infty} p(t) e^{-at} dt = \vartheta(a) < 1,$$

falls $a > 0$ ist.

²²⁾ Doetsch, S. 161, Satz IV b.
Mathematische Annalen. 118.

Für $a = 0$ würde aus

$$\int_0^{\infty} p(t) e^{-\xi t} dt = 1$$

wegen

$$\int_0^{\infty} p(t) dt = 1$$

folgen

$$\int_0^{\infty} p(t) \{1 - \cos \xi t\} dt = 0$$

oder $p(t) \{1 - \cos \xi t\} = 0$ bis auf eine Menge vom Maße Null. Da aber sicher $p(t)$ nicht äquivalent zu Null ist, hat diese Gleichung nur die Lösung $\xi = 0$.

b) Die Transformierte von $\varphi(t)$.

Um nun in der Tat $\frac{b(z)}{1-b(z)}$ als Transformierte von $\varphi(t)$ in der unteren Halbebene zu erkennen, könnten wir Gebrauch von einem Theorem von Paley und Wiener²³⁾ machen. Ersichtlich ist nämlich bei $a > 0$ $\varphi(t) e^{-a(t-k)}$ lösender Kern zu $p(t-k) e^{-a(t-k)}$. Die Fouriersche Transformierte von $p(t) e^{-at}$ ist aber $b(z - ia)$ und daher in der unteren Halbebene einschließlich der reellen Achse regulär und nach Satz 6 von 1 verschieden. Nach dem genannten Theorem folgt unter diesen Voraussetzungen dann die Konvergenz von $\int_0^{\infty} |\varphi(t)| e^{-at} dt$. Haben wir aber erst die Konvergenz von $\int_0^{\infty} |\varphi(t)| e^{-at} dt$, so folgt hieraus die absolute Konvergenz der Fourierschen Transformaten von $\varphi(t)$ in der unteren Halbebene und damit nach dem Faltungssatz sofort die Beziehung (25). Der Beweis des Theorems von Paley und Wiener stützt sich jedoch auf sehr weitgehende Sätze über die Fouriertransformation. Wir wollen daher einen einfacheren Beweis für die Konvergenz von $\int_0^{\infty} |\varphi(t)| e^{-at} dt$ geben.

Satz 7. Ist $\varphi(t)$ Lösung der Integralgleichung (5) mit einem $p(t)$ gemäß § 1, dann ist für jedes $a > 0$ $\int_0^{\infty} |\varphi(t)| e^{-at} dt$ konvergent.

Beweis. Gemäß (5) ist

$$(28) \quad |\varphi(t)| e^{-at} \leq p(t) e^{-at} + \int_0^t |\varphi(k)| e^{-ak} p(t-k) e^{-a(t-k)} dk,$$

²³⁾ Paley u. Wiener, S. 60; Theorem XVIII.

also mit einer durch (6) gerechtfertigten Vertauschung der Integrationsreihenfolge

$$\int_0^M |\varphi(t)| e^{-at} dt \leq \int_0^M p(t) e^{-at} + \int_0^M dk \int_k^M |\varphi(k)| e^{-ak} p(t-k) e^{-a(t-k)} dt$$

oder mit einer in (27) eingeführten Bezeichnungswise

$$\int_0^M |\varphi(t)| e^{-at} dt \leq \vartheta(a) + \int_0^M |\varphi(k)| e^{-ak} \vartheta(a) dk.$$

Es ist also

$$\int_0^M |\varphi(t)| e^{-at} dt \leq \frac{\vartheta(a)}{1 - \vartheta(a)}$$

oder

$$\int_0^\infty |\varphi(t)| e^{-at} dt \leq \frac{\vartheta(a)}{1 - \vartheta(a)}.$$

Damit haben wir gezeigt, daß jede Lösung $\varphi(t)$ von (5) in der unteren Halbebene die Transformierte $f(z) = \frac{\vartheta(z)}{1 - \vartheta(z)}$ hat. Hieraus folgt aber²⁴⁾, daß alle $\varphi(t)$ untereinander und insbesondere zu $\psi(t)$ äquivalent sind. Da nun aber die rechte Seite von (5) sich bei Ersetzung von $\psi(t)$ durch eine äquivalente nicht ändern würde, folgt schließlich, daß $\psi(t)$ überhaupt die einzige Lösung von (5) ist. Aber auch (4) kann nicht mehrere Lösungen haben. Sind nämlich $\varphi_1(t)$ und $\varphi_2(t)$ zwei Lösungen von (4), dann ist $\psi(t) + \varphi_1(t) - \varphi_2(t)$ Lösung von (5).

Wir fassen zusammen zu

Satz 8. Die in Satz 4 genannte Funktion $\psi(t)$ ist die einzige Lösung von (5). (9) ist die einzige Lösung von (4).

Wir können den Satz 7 nunmehr noch ergänzen durch den

Satz 7a. Gilt zusätzlich zu den Voraussetzungen des Satzes 7 noch $p(t) < M$ für alle t , so ist sogar $\lim_{t \rightarrow \infty} \psi(t) e^{-at} = 0$ für jedes $a > 0$.

Beweis. Gemäß (8) ist $\psi(t) = p(t) + r(t)$ mit nichtnegativem und stetigem $r(t)$; es genügt also, $\lim_{t \rightarrow \infty} r(t) e^{-at} = 0$ zu beweisen. In der Tat ist nach (5)

$$r(t) e^{-at} = \int_0^t \psi(k) e^{-ak} p(t-k) e^{-a(t-k)} dk,$$

²⁴⁾ Doetsch, Teil I, Kap. 3, § 7.

oder, wenn hier $\psi(t) = p(t) + r(t)$ eingesetzt wird,

$$(29) \quad r(t) e^{-at} = \int_0^t r(k) e^{-ak} p(t-k) e^{-a(t-k)} dk + e^{-at} p_2(t).$$

Ist nun $r(t) e^{-at}$ bei jedem $a > 0$ für alle t beschränkt, so ist die Behauptung des Satzes klar.

Nehmen wir also im Gegenteil an, es wäre für ein bestimmtes $a > 0$ $r(t) e^{-at}$ nicht beschränkt, dann gibt es eine Folge $t_1 < t_2 < \dots$ mit $\lim_{n \rightarrow \infty} t_n = \infty$ und $r(t_n) e^{-at_n} = \max_{0 \leq t \leq t_n} r(t) e^{-at}$, womit auch gelten muß

$$(30) \quad \lim_{n \rightarrow \infty} r(t_n) e^{-at_n} = \infty.$$

Es folgt dann aus (29)

$$\begin{aligned} r(t_n) e^{-at_n} &\leq r(t_n) e^{-at_n} \int_0^{t_n} p(t_n - k) e^{-a(t_n - k)} dk + e^{-at_n} p_2(t_n) \\ &\leq r(t_n) e^{-at_n} \theta(a) + e^{-at_n} p_2(t_n), \end{aligned}$$

woraus sich mit Hilfe von (7) ergibt

$$0 \leq r(t_n) e^{-at_n} \leq \frac{e^{-at_n} \cdot M^2 \cdot t_n}{1 - \theta(a)}$$

oder

$$\lim_{n \rightarrow \infty} r(t_n) e^{-at_n} = 0,$$

was einen Widerspruch zu (30) darstellt.

c) Die Umkehrtransformation.

Um aus $f(z)$ rückwärts $\psi(t)$ erhalten zu können, würde es bekanntlich genügen, wenn $\psi(t)$ stetig und $f(z)$ längs einer in der unteren Halbebene verlaufenden Parallelen zur reellen Achse absolut integrierbar²⁵⁾ wäre. Das ist aber im allgemeinen nicht erfüllt. Wir können jedoch die angegebenen Bedingungen sofort erhalten, wenn wir nicht $\psi(t)$, sondern seinen stetigen Bestandteil $r(t) = \psi(t) - p(t)$ betrachten und über $p(t)$ nur noch voraussetzen, daß es quadratintegrierbar ist²⁶⁾.

Aus den Transformaten von $\psi(t)$ und $p(t)$ berechnen wir zunächst die Transformaten von $r(t)$ zu

$$(31) \quad \varrho(z) = \frac{b(z)}{1 - b(z)} - b(z) = \frac{b^2(z)}{1 - b(z)}.$$

²⁵⁾ Doetsch, S. 107, Satz 6.

²⁶⁾ Hieraus folgt zwar nach Plancherel die Fourier-Darstellbarkeit von $p(t)$ im Mittel; da wir eine Darstellung für jedes t suchen, machen wir keinen Gebrauch von diesem Satz.

Damit nun für ein $a > 0$ das Integral $\int_{-ia-\infty}^{-ia+\infty} |\varrho(z)| dz$ konvergiert, ist auf Grund von (27) notwendig, daß $\int_{-ia-\infty}^{-ia+\infty} |b(z)|^2 dz$ konvergiert. Das ist aber für quadratintegrierbares $p(t)$ sichergestellt durch einen Satz von Plancherel²⁷⁾, der auf unseren Fall angewendet bei beliebigem $a \geq 0$ lautet

$$(32) \quad \int_{-ia-\infty}^{-ia+\infty} |b(z)|^2 dz = \int_{-\infty}^{+\infty} |b(\xi - ia)|^2 d\xi = 2\pi \int_0^{\infty} p^2(t) e^{-2at} dt,$$

da ja $b(z - ia)$ die Fourier-Transformierte der absolut und quadratintegrierbaren Funktion $p(t) e^{-at}$ ist.

$r(t)$ ist also stetig und in jedem endlichen Intervall beschränkt, seine Fouriersche Transformierte konvergiert überall in der unteren Halbebene und ist dort längs jeder Parallelen zur reellen Achse absolut integrierbar; dann gilt aber²⁸⁾

Satz 9. Für jedes $a > 0$ ist

$$(33) \quad r(t) = \frac{1}{2\pi} \int_{-ia-\infty}^{-ia+\infty} \frac{b^2(z)}{1-b(z)} e^{iat} dz.$$

§ 5.

Verschiebung des Integrationsweges.

Wie wir an der zu (33) gleichwertigen Formel

$$(34) \quad r(t) e^{-at} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{b^2(\xi - ia)}{1-b(\xi - ia)} e^{i\xi t} d\xi$$

sehen, ist der Grenzübergang $a \rightarrow 0$ nötig, um auf Grund des Riemannschen Lemmas²⁹⁾ Aussagen über das Verhalten von $r(t)$ bei großen t machen zu können. Die damit für (33) geforderte Verschiebung des Integrationsweges auf die reelle Achse als Rand der Regularitätshalbebene ist bekanntlich³⁰⁾ durchführbar, wenn nach Abspaltung leicht zu übersehender singulärer Teile von $\varrho(z)$ der Rest bis einschließlich der reellen Achse stetig ist und sonst noch geeignete Abschätzungen gelten. Wir wollen jedoch einen anderen Weg ein-

²⁷⁾ Sog. Parseval'sche Gleichung. Bochner, S. 172, Satz 51. Die Normierung ist dort etwas anders.

²⁸⁾ Doetsch, S. 107, Satz 6.

²⁹⁾ Doetsch, S. 50, Satz 3.

³⁰⁾ Vgl. Doetsch, S. 269, Satz 4.

schlagen, der diese Verschiebung des Integrationsweges unter etwas allgemeineren Bedingungen³¹⁾ als den üblichen zuläßt: (34) wird in Real- und Imaginärteil gespalten und der Grenzübergang $a \rightarrow 0$ der Teile einzeln untersucht. Es zeigt sich, daß der eine Bestandteil gegen die Verschiebung des Integrationsweges unempfindlicher ist als der andere und damit die komplexe Formel. Für das Erneuerungsproblem hat dies zur Folge, daß z. B. $P(t) = \frac{1}{(1+t)^2}$ noch zugelassen bleibt, während es sonst ausgeschlossen werden müßte.

Zur Durchführung des angegebenen Programms setzen wir zunächst bei $z = \xi - ia$

$$b(\xi - ia) = b_1(\xi, a) - ib_2(\xi, a)$$

mit reellen $b_1(\xi, a)$, wobei

$$(35) \quad \left\{ \begin{array}{l} b_1(\xi, a) = \int_0^{\infty} p(t) e^{-at} \cos \xi t dt \quad \text{eine gerade} \\ \text{und } b_2(\xi, a) = \int_0^{\infty} p(t) e^{-at} \sin \xi t dt \quad \text{eine ungerade Funktion von } \xi \text{ ist.} \end{array} \right.$$

Nach (31) ist dann

$$\varrho(\xi - ia) = \varrho_1(\xi, a) - i\varrho_2(\xi, a)$$

mit

$$(36) \quad \left\{ \begin{array}{l} \varrho_1(\xi, a) = \frac{b_1^2 - b_2^2 - b_1^2 - b_2^2}{(1-b_1)^2 + b_2^2} \\ \text{und } \varrho_2(\xi, a) = \frac{2b_1 b_2 - b_1^2 b_2 - b_2^3}{(1-b_1)^2 + b_2^2}. \end{array} \right.$$

Damit wird nun nach (34)

$$\begin{aligned} 2\pi r(t) e^{-at} &= \int_{-\infty}^{+\infty} (\varrho_1 - i\varrho_2) (\cos \xi t + i \sin \xi t) d\xi \\ &= \int_{-\infty}^{+\infty} (\varrho_1 \cos \xi t + \varrho_2 \sin \xi t) d\xi + i \int_{-\infty}^{+\infty} (\varrho_1 \sin \xi t - \varrho_2 \cos \xi t) d\xi, \end{aligned}$$

wobei das zweite Integral verschwindet, da der Integrand ungerade ist. Für positive t ist daher wegen $r(-t) = 0$:

$$\begin{aligned} 2\pi r(t) e^{-at} &= 2 \int_0^{\infty} (\varrho_1 \cos \xi t + \varrho_2 \sin \xi t) d\xi \\ 0 &= 2 \int_0^{\infty} (\varrho_1 \cos \xi t - \varrho_2 \sin \xi t) d\xi, \end{aligned}$$

³¹⁾ Liegt die Singularität bei $z = 0$, so muß dort $|\text{Restfunktion}| \leq \frac{\text{Const.}}{|\text{Realteil von } z|}$ sein.

und daraus folgt durch Addition und Subtraktion

$$\frac{\pi}{2} r(t) e^{-at} = \int_0^{\infty} \varrho_1 \cos \xi t d\xi$$

und $\frac{\pi}{2} r(t) e^{-at} = \int_0^{\infty} \varrho_2 \sin \xi t d\xi,$

woraus sich mit Hilfe von (36) schließlich ergibt:

Für $t > 0$ ist

$$(37a) \quad r(t) = \lim_{a \rightarrow +0} \frac{2}{\pi} \int_0^{\infty} \frac{b_1^2 - b_2^2 - b_1^2 b_2^2}{(1 - b_1)^2 + b_2^2} \cos \xi t d\xi,$$

$$(37b) \quad r(t) = \lim_{a \rightarrow +0} \frac{2}{\pi} \int_0^{\infty} \frac{2b_1 b_2 - b_1^2 b_2 - b_2^2}{(1 - b_1)^2 + b_2^2} \sin \xi t d\xi \text{ mit } b_r = b_r(\xi, a) \text{ bei } a > 0.$$

Hier wollen wir nun versuchen, den Grenzübergang unter dem Integral durchzuführen.

Zunächst überzeugen wir uns leicht, daß dies bei (37a) nicht zulässig ist.

So liefert z. B. $p(t) = e^{-t}$ die Transformierte $b(z) = \frac{1}{1+iz}$ und hieraus $\varrho(z) = \frac{1}{iz(1+iz)}$. Also ist $\varrho_1(\xi, a) = \frac{a+a^2-\xi^2}{(\xi^2+a^2)(1+a^2+\xi^2)}$, $\varrho_1(\xi, 0) = \frac{-1}{1+\xi^2}$. Ließe sich nun in (37a) der Grenzübergang unter dem Integralzeichen durchführen, so wäre nach Laplace

$$r(t) = \frac{-2}{\pi} \int_0^{\infty} \frac{\cos \xi t}{1+\xi^2} d\xi = -e^{-t},$$

was aber $r(t) > 0$ widerspricht.

Wir kommen nun zur Umformung der Relation (37b). Aus (35) gewinnen wir zunächst durch Differenzieren nach a

$$\frac{\partial b_1(\xi, a)}{\partial a} = - \int_0^{\infty} x p(x) e^{-ax} \cos \xi x dx \quad \text{und} \quad \frac{\partial b_2(\xi, a)}{\partial a} = - \int_0^{\infty} x p(x) e^{-ax} \sin \xi x dx$$

und hieraus

$$\left| \frac{\partial b_r(\xi, a)}{\partial a} \right| \leq \int_0^{\infty} x p(x) dx = \bar{x}.$$

Also wird

$$|b_r(\xi, a) - b_r(\xi, 0)| \leq a \bar{x}, \quad r = 1, 2.$$

Da somit $b, (\xi, a)$ für $a \rightarrow 0$ gleichmäßig in ξ gegen $b, (\xi, 0)$ strebt und andererseits $\lim_{\xi \rightarrow \pm\infty} b, (\xi, 0) = 0$ ist, läßt sich jedenfalls ein Bereich $0 \leq a \leq a_1$, $|\xi| \geq B_1$ so abgrenzen, daß der Nenner von (37b) größer als eine feste positive Schranke bleibt.

Betreffs des Zählers beachten wir, daß nach der Parsevalschen Gleichung (32)

$$\lim_{a \rightarrow +0} \int_{-\infty}^{+\infty} |b(\xi, a)|^2 d\xi = \int_{-\infty}^{+\infty} |b(\xi, 0)|^2 d\xi$$

ist, da ja die rechte Seite von (32) offenbar bei $a \rightarrow 0$ Konvergenz zeigt und (32) auch bei $a = 0$ gilt. Für endliche Intervallgrenzen ist aber die entsprechende Relation wegen der soeben festgestellten gleichmäßigen Konvergenz von $b, (\xi, a)$ gegen $b, (\xi, 0)$ selbstverständlich, so daß unter Beachtung der Geradheit von $|b(\xi, a)|^2$ für jedes x gilt:

$$\lim_{a \rightarrow +0} \int_x^\infty |b(\xi, a)|^2 d\xi = \int_x^\infty |b(\xi, 0)|^2 d\xi.$$

Da nun weiter wieder wegen (32) $\lim_{x \rightarrow \infty} \int_x^\infty |b(\xi, 0)|^2 d\xi = 0$ ist, folgern wir, daß es zu jedem $\varepsilon > 0$ ein positives a_2 und ein B_2 gibt, so daß für $0 \leq a \leq a_2$

$$\int_{B_2}^\infty |b(\xi, a)|^2 d\xi < \varepsilon$$

wird.

Dies gilt natürlich erst recht, wenn wir $|b(\xi, a)|^2$ durch $b_1^2(\xi, a)$, $b_2^2(\xi, a)$ oder $|b_1(\xi, a) \cdot b_2(\xi, a)|$ ersetzen, endlich aber wegen der Beschränktheit von $b(\xi, a)$ auch mit $|b_1^2(\xi, a) b_2(\xi, a)|$ und mit $|b_2(\xi, a)|^3$ als Integranden.

Zusammen mit der obigen Abschätzung des Nenners erhalten wir somit:

$$(38) \quad \left\{ \begin{array}{l} \text{Zu jedem } \varepsilon > 0 \text{ gibt es ein } B \text{ und ein } a_0 > 0 \text{ so, daß} \\ \int_B^\infty |q_2(\xi, a)| d\xi < \varepsilon \text{ wird für } 0 \leq a \leq a_0. \end{array} \right.$$

Nachdem damit das Verhalten des in (37b) stehenden Integrals für die Umgebung von $\xi = \infty$ geklärt ist, wenden wir uns jetzt der Umgebung von $\xi = 0$ zu. Wir zeigen dazu zunächst den

Hilfssatz. Setzt man $\int_0^\infty x p(x) e^{-ax} dx = \bar{x}(a)$, so gilt gleichmäßig in a

$$\lim_{\xi \rightarrow +0} \frac{b_2(\xi, a)}{\xi} = \bar{x}(a).$$

Beweis. Es ist nach (35)

$$\bar{x}(a) - \frac{b_2(\xi, a)}{\xi} = \int_0^{\infty} x p(x) e^{-ax} \left\{ 1 - \frac{\sin \xi x}{\xi x} \right\} dx > 0.$$

Sei nun ein $\varepsilon > 0$ gegeben, dann wählen wir eine Zahl C so groß, daß

$$\int_C^{\infty} x p(x) dx < \varepsilon$$

wird. Weiter wählen wir $y_0 > 0$ so klein, daß für $0 \leq y \leq y_0$ $1 \geq \frac{\sin y}{y} \geq 1 - \varepsilon$ gilt. Es ist dann für $\xi < \frac{y_0}{C}$

$$\begin{aligned} 0 < \bar{x}(a) - \frac{b_2(\xi, a)}{\xi} &= \int_0^C x p(x) e^{-ax} \left\{ 1 - \frac{\sin \xi x}{\xi x} \right\} dx + \int_C^{\infty} x p(x) e^{-ax} \left\{ 1 - \frac{\sin \xi x}{\xi x} \right\} dx \\ &\leq \int_0^C x p(x) \varepsilon dx + \int_C^{\infty} x p(x) \cdot 1 \cdot 2 dx \leq \varepsilon (\bar{x} + 2). \end{aligned}$$

Aus dem Hilfssatz ziehen wir nun vermöge der selbstverständlichen Relation $\lim_{a \rightarrow +0} \bar{x}(a) = \bar{x}$ sofort die Folgerung:

$$(39) \quad \frac{b_2(\xi, a)}{\xi} \text{ ist stetig in } \xi = +0, a = +0 \text{ mit dem Werte } \bar{x}.$$

Es hat also $b_2(\xi, a)$ für ein gewisses Intervall $0 < \xi < \xi_0$ sicher keine Nullstelle und wir können daher für dieses Intervall das in (37b) stehende Produkt $\varrho_2(\xi, a) \cdot \sin \xi t$ umformen zu

$$\varrho_2(\xi, a) \sin \xi t = (2b_1 - b_1^2 - b_2^2) \cdot \frac{b_2^2}{(1-b_1)^2 + b_2^2} \cdot \frac{\xi}{b_2(\xi, a)} \cdot \frac{\sin \xi t}{\xi}.$$

Für eine (ξ, a) -Umgebung von $\xi = a = +0$ sind nun sicher der erste und nach (39) auch der dritte Faktor beschränkt; halten wir ein bestimmtes t fest, so zeigt auch der vierte Faktor diese Eigenschaft. Solange schließlich $\xi \neq 0$ bleibt, kann nach Satz 6 nicht gleichzeitig $b_1 = 1$ und $b_2 = 0$ sein, so daß der zweite Faktor immer bestimmt ist mit einem Werte kleiner als 1. Also ist $\varrho_2(\xi, a) \cdot \sin \xi t$ bei festem t in einer gewissen Umgebung von $\xi = a = +0$ beschränkt. Zu jedem $\varepsilon > 0$ gibt es also bei festem t ein $\beta > 0$ so klein, daß

$$(40) \quad \int_0^{\beta} |\varrho_2(\xi, a) \sin \xi t| d\xi < \varepsilon \text{ wird bei } 0 \leq a \leq a_0 \text{ mit } a_0 > 0.$$

(38) und (40) liefern zusammen mit (37b):

$$(41) \quad \left\{ \begin{array}{l} \text{Bei festem } t \text{ gibt es zu jedem } \varepsilon > 0 \text{ ein } a_0 > 0, \text{ ein } B \text{ und ein } \beta > 0 \\ \text{so, daß} \\ \left| r(t) - \frac{2}{\pi} \int_{\beta}^B \frac{2b_1 b_2 - b_1^2 b_2 - b_2^2}{(1-b_1)^2 + b_2^2} \sin \xi t d\xi \right| < \varepsilon \\ \text{wird bei } 0 < a \leq a_0, \text{ nicht dagegen zunächst für } a = 0. \end{array} \right.$$

In $\beta \leq \xi \leq B$ wird nun nach Satz 6 der Nenner $(1-b_1)^2 + b_2^2$ nirgends Null. Da nun $b_r(\xi, a)$ bei $a \rightarrow 0$ stetig in $b_r(\xi, 0)$ übergehen, gilt (41) auch für $a = 0$. Da auch (38) und (40) für $a = 0$ gelten, folgt

$$\left| r(t) - \frac{2}{\pi} \int_0^{\infty} \frac{2b_1(\xi, 0)b_2(\xi, 0) - b_1^2(\xi, 0)b_2(\xi, 0) - b_2^3(\xi, 0)}{[1-b_1(\xi, 0)]^2 + b_2^2(\xi, 0)} \sin \xi t d\xi \right| < \varepsilon,$$

und da dies für jedes ε gilt, bleibt

Satz 10. Für $t > 0$ ist

$$r(t) = \frac{2}{\pi} \int_0^{\infty} \frac{2b_1 b_2 - b_1^2 b_2 - b_2^3}{(1-b_1)^2 + b_2^2} \sin \xi t d\xi \quad \text{mit} \quad b_r = b_r(\xi, 0).$$

Damit ist in der Tat die Vertauschung von Integral und Grenzübergang in (37b) als statthaft erkannt.

§ 6.

Das Verhalten für große t .

Es sei ein $\delta > 0$ beliebig gewählt, dann ist nach Satz 10

$$r(t) = \frac{2}{\pi} \int_0^{\delta} \varrho_2(\xi, 0) \sin \xi t d\xi + \frac{2}{\pi} \int_{\delta}^{\infty} \varrho_2(\xi, 0) \sin \xi t d\xi.$$

Da nun $b(\xi, 0)$ eine stetige Funktion von ξ und für $\xi > 0$ sicher $(1-b_1)^2 + b_2^2 > 0$ ist, muß für jedes B

$$\lim_{t \rightarrow \infty} \int_{\delta}^B \varrho_2(\xi, 0) \sin \xi t d\xi = 0$$

sein. In Verbindung mit (38) folgt daraus aber

$$\lim_{t \rightarrow \infty} \int_0^{\infty} \varrho_2(\xi, 0) \sin \xi t d\xi = 0.$$

Es genügt somit, $\lim_{t \rightarrow \infty} \frac{2}{\pi} \int_0^{\delta} \varrho_2(\xi, 0) \sin \xi t d\xi$ zu betrachten, wobei wir δ sehr klein wählen können. Nun können wir schreiben

$$\varrho_2(\xi, 0) = \frac{2b_1b_2 - b_1^2b_2 - b_2^3}{(1-b_1)^2 + b_2^2} = \frac{b_2}{(1-b_1)^2 + b_2^2} - b_2, \quad b_1 = b_1(\xi, 0).$$

Da $b_2(\xi, 0)$ stetig ist, wird sicher $\lim_{t \rightarrow \infty} \frac{2}{\pi} \int_0^{\delta} b_2(\xi, 0) \sin \xi t d\xi = 0$, so daß bleibt

$$(42) \quad \lim_{t \rightarrow \infty} r(t) = \lim_{t \rightarrow \infty} \frac{2}{\pi} \int_0^{\delta} \frac{b_2}{(1-b_1)^2 + b_2^2} \sin \xi t d\xi,$$

wobei das Zeichen \lim auf beiden Seiten der Gleichung durch $\underline{\lim}$ oder durch $\overline{\lim}$ ersetzt zu denken ist.

Wir führen jetzt zweckmäßig die Transformierte $B(z)$ von $P(t)$ ein und setzen

$$(43) \quad \left\{ \begin{array}{l} B(\xi - i0) = B(\xi) = B_1(\xi) - iB_2(\xi) \\ \text{mit } B_1(\xi) = \int_0^{\infty} P(t) \cos \xi t dt \\ \text{und } B_2(\xi) = \int_0^{\infty} P(t) \sin \xi t dt; \end{array} \right.$$

dann gilt wegen der durch partielle Integration leicht zu erhaltenden Beziehung

$$(44) \quad b(z) = 1 - izB(z) \quad \text{mit } B(0) = \bar{x},$$

daß

$$b_1(\xi, 0) = 1 - \xi \cdot B_2(\xi)$$

$$\text{und} \quad b_2(\xi, 0) = \xi \cdot B_1(\xi)$$

ist. In (42) eingesetzt liefert dies

$$(45) \quad \lim_{t \rightarrow \infty} r(t) = \lim_{t \rightarrow \infty} \frac{2}{\pi} \int_0^{\delta} \frac{B_1(\xi)}{B_1^2(\xi) + B_2^2(\xi)} \cdot \frac{\sin \xi t}{\xi} d\xi.$$

Damit haben wir ein Dirichletsches Integral ³²⁾ erhalten, wo die Funktion $\gamma(\xi) = \frac{B_1(\xi)}{B_1^2(\xi) + B_2^2(\xi)}$ stetig ist mit $\lim_{\xi \rightarrow 0} \gamma(\xi) = \frac{\bar{x}}{\bar{x}^2 + 0^2} = \frac{1}{\bar{x}}$.

Um aus (45) demgemäß $\lim_{t \rightarrow \infty} r(t) = \frac{2}{\pi} \cdot \frac{\pi}{2} \cdot \frac{1}{\bar{x}} = \frac{1}{\bar{x}}$ schließen zu können, genügt bekanntlich im allgemeinen die Stetigkeit von $\gamma(\xi)$ nicht. Dagegen

³²⁾ Vgl. etwa Bochner, Kap. II, § 7.

ist es ohne weitere Einschränkung möglich, die Stabilisierung im Mittel zu beweisen. Setzen wir nämlich bei festem δ

$$(46) \quad r_1(t) = \frac{2}{\pi} \int_0^{\delta} \gamma(\xi) \frac{\sin \xi t}{\xi} d\xi,$$

so ist nach (45) $\lim_{t \rightarrow \infty} [r(t) - r_1(t)] = 0$ und damit

$$(47) \quad \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t r(x) dx = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t r_1(x) dx.$$

Nun berechnen wir aber leicht

$$\frac{1}{t} \int_0^t r_1(x) dx = \frac{2}{\pi t} \int_0^{\delta} \gamma(\xi) \frac{1 - \cos \xi t}{\xi^2} d\xi = \frac{2}{\pi} \int_0^{\delta} \gamma(\xi) \frac{\sin^2 \frac{\xi t}{2}}{\left(\frac{\xi t}{2}\right)^2} d\left(\frac{\xi t}{2}\right).$$

Das ist nun ein Féjersches Integral³³⁾, bei dem wir wegen der Stetigkeit von $\gamma(\xi)$ sofort zur Grenze $t \rightarrow \infty$ übergehen können mit

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t r_1(x) dx = \lim_{\xi \rightarrow +0} \gamma(\xi) = \frac{1}{2}.$$

Für die Erneuerungsfunktion erhalten wir hieraus wegen (8) und (47)

$$\frac{1}{t} \Psi(t) = \frac{1}{t} \int_0^t \psi(x) dx = \frac{1 - P(t)}{t} + \frac{1}{t} \int_0^t r(x) dx,$$

und daher $\lim_{t \rightarrow \infty} \frac{1}{t} \Psi(t) = \frac{1}{2}.$

Wir fassen zusammen zu

Satz 11. Erfüllt die Ausscheidungsfunktion die Bedingungen: $\int_0^{\infty} t p(t) dt = \bar{x}$ ist endlich, und $\int_0^{\infty} p^2(t) dt$ ist endlich, dann zeigt der Erneuerungsvorgang Stabilisierung im Mittel mit dem Werte $\frac{1}{2}.$

Um nun auch über die eigentliche Stabilisierung Aussagen machen zu können, wollen wir die an $p(t)$ und $P(t)$ gestellten Voraussetzungen noch ein wenig verschärfen. Und zwar sei im folgenden vorausgesetzt, daß nicht nur $\int_0^{\infty} P(t) dt$, sondern sogar noch $\int_1^{\infty} P(t) \ln t dt$ konvergiert. Wir bemerken hierzu:

³³⁾ Vgl. etwa Bochner, S. 22.

Wenn sich $P(t)$ für große t wie $\frac{c}{t^n}$ verhält, so mußte bisher $n > 1$ gefordert werden. Diese Bedingung ist jedoch auch für die angegebene verschärfte Voraussetzung hinreichend; wegen

$$\int_1^{\infty} \frac{\ln^k x}{x^n} dx = \int_0^{\infty} y^k e^{-(n-1)y} dy = \frac{\Gamma(k+1)}{(n-1)^{k+1}} = \text{endlich} \quad \text{für } k \geq 0, n > 1$$

bleiben nämlich wie bisher alle $P(t)$ in die Betrachtung eingeschlossen, die sich wie $c \cdot \frac{\ln^k t}{t^n}$ bei $n > 1$ verhalten.

Es sei nun in der Tat die Konvergenz von $\int_0^{\infty} P(t) \ln t dt$ vorausgesetzt, dann betrachten wir zunächst bei $0 < \delta_1 < \delta$ die Funktion

$$u(\delta_1, \delta) = \int_{\delta_1}^{\delta} \frac{\bar{x} - B_1(\xi)}{\xi} d\xi = \int_{\xi=\delta_1}^{\delta} \int_{x=0}^{\infty} P(x) \frac{1 - \cos \xi x}{\xi} dx d\xi.$$

Wegen der Beschränktheit von $\left| \frac{1 - \cos \xi x}{\xi} \right| < \frac{2}{\delta_1}$ mit $\delta_1 > 0$ und der Konvergenz von $\int_0^{\infty} P(t) dt$ können die Integrationen vertauscht werden, und damit ist

$$u(\delta_1, \delta) = \int_{x=0}^{\infty} P(x) \int_{\delta_1}^{\delta} \frac{1 - \cos \xi x}{\xi} d\xi dx = \int_{x=0}^{\infty} P(x) \int_{\delta_1 x}^{\delta x} \frac{1 - \cos t}{t} dt dx.$$

Da $\frac{1 - \cos t}{t} \geq 0$ für alle t ist, wird

$$u(\delta_1, \delta) \leq \int_{x=0}^{\infty} P(x) \int_0^{\delta x} \frac{1 - \cos t}{t} dt dx$$

und damit auch

$$\int_0^{\delta} \frac{\bar{x} - B_1(\xi)}{\xi} d\xi = \lim_{\delta_1 \rightarrow +0} u(\delta_1, \delta) \leq \int_{x=0}^{\infty} P(x) \int_0^{\delta x} \frac{1 - \cos t}{t} dt dx.$$

Nun gilt für alle t die Abschätzung $\frac{1 - \cos t}{t} < \frac{3}{1+t}$, und damit wird

$$(48) \quad \int_0^{\delta} \frac{\bar{x} - B_1(\xi)}{\xi} d\xi < 3 \int_0^{\infty} P(x) \ln(1 + \delta x) dx,$$

was nach Voraussetzung sicher konvergiert.

Es sei nunmehr ein ε mit $0 < \varepsilon < 1$ beliebig gegeben, dann wählen wir ein $B > 2$ so groß, daß $\int_B^\infty P(t) \ln t \, dt < \varepsilon$ wird. Dann ist für $\delta < \frac{\varepsilon}{B}$

$$\begin{aligned} \int_0^\infty P(x) \ln(1 + \delta x) \, dx &= \int_0^B P(x) \ln(1 + \delta x) \, dx + \int_B^\infty P(x) \ln(1 + \delta x) \, dx \\ &< \varepsilon \int_0^B P(x) \, dx + \int_B^\infty P(x) \ln x \, dx < \varepsilon(1 + \bar{x}), \end{aligned}$$

und damit

$$\lim_{\delta \rightarrow +0} \int_0^\infty P(x) \ln(1 + \delta x) \, dx = 0.$$

Aus (48) folgt nun sofort

$$(49) \quad \lim_{\delta \rightarrow +0} \int_0^\delta \frac{\bar{x} - B_1(\xi)}{\xi} \, d\xi = 0.$$

Wörtlich genau so wie oben gewinnen wir die Abschätzung

$$\begin{aligned} v(\delta) &= \int_0^\delta \int_0^\infty \int_0^\infty P(x) P(y) \frac{1 - \cos \xi(x-y)}{\xi} \, dx \, dy \, d\xi \\ &\leq \int_0^\infty \int_0^\infty P(x) P(y) \int_0^\delta \frac{1 - \cos \xi(x-y)}{\xi} \, d\xi \, dx \, dy \\ &< 3 \int_0^\infty \int_0^\infty P(x) P(y) \ln(1 + \delta |x-y|) \, dx \, dy, \end{aligned}$$

oder unter Beachtung der Symmetrie des Integranden und bei $\begin{cases} x' = x-y \\ y' = y \end{cases}$

$$\begin{aligned} v(\delta) &< 6 \int_0^\infty \int_0^\infty P(x' + y') P(y') \ln(1 + \delta x') \, dx' \, dy' \\ &\leq 6 \int_0^\infty \int_0^\infty P(x') P(y') \ln(1 + \delta x') \, dx' \, dy' = 6 \bar{x} \int_0^\infty P(x') \ln(1 + \delta x') \, dx', \end{aligned}$$

und damit

$$\lim_{\delta \rightarrow +0} v(\delta) = \lim_{\delta \rightarrow +0} \int_0^\delta \int_0^\infty \int_0^\infty P(x) P(y) \frac{1 - \cos \xi(x-y)}{\xi} \, dx \, dy \, d\xi = 0.$$

Ebenso ist natürlich auch

$$\lim_{\delta \rightarrow +0} \int_0^\delta \int_0^\infty \int_0^\infty P(x) P(y) \frac{1 - \cos \xi(x+y)}{\xi} \, dx \, dy \, d\xi = 0,$$

und damit wegen

$$\begin{aligned} B_2^2(\xi) &= \iint_0^\infty P(x) P(y) \sin \xi x \sin \xi y dx dy \\ &= \frac{1}{2} \iint_0^\infty P(x) P(y) [1 - \cos \xi(x+y)] dx dy + \\ &\quad - \frac{1}{2} \iint_0^\infty P(x) P(y) [1 - \cos \xi(x-y)] dx dy \end{aligned}$$

endlich

$$(50) \quad \lim_{\delta \rightarrow 0} \int_0^\delta \frac{B_2^2(\xi)}{\xi} d\xi = 0.$$

Setzen wir nun in (46)

$$\frac{B_1(\xi)}{B_1^2(\xi) + B_2^2(\xi)} = \frac{1}{\bar{x}} + \frac{\bar{x} - B_1(\xi)}{\bar{x} \cdot B_1(\xi)} - \frac{B_2^2(\xi)}{B_1(\xi) \cdot [B_1^2(\xi) + B_2^2(\xi)]},$$

so haben wir

$$\begin{aligned} \frac{\pi}{2} r_1(t) &= \frac{1}{\bar{x}} \int_0^\delta \frac{\sin \xi t}{\xi} d\xi + \int_0^\delta \frac{\bar{x} - B_1(\xi)}{\xi} \cdot \frac{\sin \xi t}{\bar{x} \cdot B_1(\xi)} d\xi + \\ &\quad - \int_0^\delta \frac{B_2^2(\xi)}{\xi} \cdot \frac{\sin \xi t}{B_1(\xi) \cdot [B_1^2(\xi) + B_2^2(\xi)]} d\xi. \end{aligned}$$

Ist dann ein $\varepsilon > 0$ beliebig gegeben, so können wir uns wegen $\lim_{\xi \rightarrow 0} B_1(\xi) = \bar{x}$ mit (49) und (50) δ als so klein vorstellen, daß die letzten beiden Summanden zusammen ε nicht übersteigen. Es ist dann nach (45)

$$\begin{aligned} \frac{\pi}{2} \overline{\lim}_{t \rightarrow \infty} r(t) &\leq \frac{1}{\bar{x}} \cdot \frac{\pi}{2} + \varepsilon, \\ \frac{\pi}{2} \underline{\lim}_{t \rightarrow \infty} r(t) &\geq \frac{1}{\bar{x}} \cdot \frac{\pi}{2} - \varepsilon, \end{aligned}$$

und da ε beliebig klein gewählt werden kann,

$$(51) \quad \lim_{t \rightarrow \infty} r(t) = \frac{1}{\bar{x}}.$$

Nun ist $\psi(t) = p(t) + r(t)$. Wir sahen schon in Satz 5, daß für die eigentliche Stabilisierung $\lim_{t \rightarrow \infty} \hat{p}(t) = 0$ notwendig war, wobei $\hat{p}(t)$ äquivalent zu $p(t)$ ist. Aus (51) folgt nun unter den gemachten Einschränkungen, daß diese Bedingung auch hinreichend ist.

Satz 12. Erfüllt die Ausscheidungsfunktion $p(t)$ die Bedingungen:

$$\int_0^{\infty} t p(t) \ln t \, dt \text{ ist endlich, und } p(t) \text{ ist quadratintegrierbar,}$$

dann ist $\psi(t) = p(t) + \frac{1}{x} + \chi(t)$ mit $\lim_{t \rightarrow \infty} \chi(t) = 0$.

Notwendig und hinreichend für eigentliche Stabilisierung des Erneuerungsvorgangs ist, daß $\lim_{t \rightarrow \infty} \hat{p}(t) = 0$ für eine passende äquivalente $\hat{p}(t)$ zu $p(t)$ gilt. Anderenfalls kopiert $\psi(t)$ schließlich getreu alle Schwankungen von $p(t)$.

§ 7.

Verallgemeinerung der Ergebnisse.

Wir waren bei der Formulierung des Erneuerungsproblems davon ausgegangen, daß zur Zeit $t = 0$ die betrachtete Gesamtheit lediglich aus 0-Jährigen besteht. Diese Voraussetzung ist selbstverständlich ziemlich willkürlich; es wird auch in allen praktischen Fällen so sein, daß bereits zur Zeit $t = 0$ eine gewisse Altersgliederung der Anfangsgesamtheit statthat. Wir stellen uns daher die Frage, ob unsere Ergebnisse gültig bleiben, wenn wir eine beliebig gegliederte Ausgangsgesamtheit zulassen. Ebenso lassen wir jetzt die Voraussetzung fallen, daß ein konstanter Bestand vorgeschrieben ist, und lassen statt dessen eine allgemeinere Bestandsentwicklung $H(t)$ zu.

Bezeichnen wir mit $F(k)$ die Anzahl der 0- bis k -jährigen Mitglieder der Anfangsgesamtheit, so ist $F(k)$ eine monoton nichtfallende Funktion, die wir als halbstetig nach links ansehen können. Bei endlichem Höchstalter ω sollen überdies keine Mitglieder vom Alter ω vorhanden sein, da dieselben ja sofort ersetzt werden müßten und daher gleich als 0-Jährige zu schreiben sind. Mit dieser Normierung wird $F(0) = 0$ und für endliches oder unendliches Höchstalter $F(\omega) = F(\omega - 0) = 1$.

Von $dF(k)$ Mitgliedern des Alters k bis $k + dk$ werden nun nach der Zeit t noch $dF(k) \frac{P(k+t)}{P(k)}$ der Gesamtheit angehören, so daß sich die Überlebenden der Anfangsgesamtheit zur Zeit t insgesamt ergeben zu

$$(52) \quad P^*(t) = \int_0^{\omega} \frac{P(k+t)}{P(k)} dF(k).$$

Wegen des monotonen Nichtsteigens von $P(t)$ ist tatsächlich $P^*(t)$ für jedes t bei beliebigem $F(k)$ definiert und ist selbst monoton nichtsteigend mit $P^*(0) = 1$. Schreiben wir mit beliebigem $x < \omega$

$$P^*(t) = \int_0^x \frac{P(k+t)}{P(k)} dF(k) + \int_x^{\omega} \frac{P(k+t)}{P(k)} dF(k),$$

so ist ersichtlich der zweite Summand wegen

$$\int_x^{\infty} \frac{P(k+t)}{P(k)} dF(k) \leq \int_x^{\infty} dF(k)$$

bei genügend großem x beliebig klein zu machen. Der erste Summand ist aber bei festem x wegen der gleichmäßigen Stetigkeit von $P(t)$ auch stetig und strebt mit wachsendem t gegen Null. Also ist überhaupt $P^*(t)$ stetig und genügt der Bedingung

$$\lim_{t \rightarrow \infty} P^*(t) = 0.$$

Stirbt nun die Anfangsgesamtheit gemäß der Funktion $P^*(t)$ ab und soll die Gesamtheit auf der Höhe $H(t)$ gehalten werden, so muß der Zugang $\Phi(t)$ in Analogie zu (3) ersichtlich der Integralgleichung

$$(53) \quad H(t) = P^*(t) + \int_0^t P(t-k) d\Phi(k)$$

genügen.

Wenn $H(t)$ in jedem endlichen Intervall beschränkt ist, so können wir wegen der für $P^*(t)$ bewiesenen Eigenschaften genau so wie im § 1 schließen, daß auch $\Phi(t)$ in jedem endlichen Intervall beschränkt ist; vorausgesetzt, daß (53) überhaupt lösbar ist. Ebenso folgern wir: Ist $H(t)$ stetig, so auch $\Phi(t)$ und umgekehrt; ist $H(t) - P^*(t)$ totalstetig, so auch $\Phi(t)$ und umgekehrt. $H(t)$ kann nicht ganz beliebig gewählt werden, da ja $d\Phi$ seiner Bedeutung gemäß positiv sein muß. Trotzdem können wir aber rein formal (53) auch in solchen Fällen als Aufgabe zulassen, wo $H(t)$ in diesem Sinne für ein Erneuerungsproblem ungeeignet ist.

Durch partielle Integration erhalten wir nun aus (53)

$$\Phi(t) = H(t) - P^*(t) + \int_0^t p(t-k) \Phi(k) dk.$$

Diese Gleichung ist aber ebenso wie (5) eindeutig lösbar, und wir erhalten unter Berücksichtigung der Tatsache, daß $\psi(t-k)$ der lösende Kern ist, sofort

Satz 13. Bei beliebiger Anfangsalterszusammensetzung und dem vorgeschriebenen Gesamtbestande $H(t)$ ist der Gesamtzugang gegeben durch

$$\Phi(t) = H(t) - P^*(t) + \int_0^t [H(k) - P^*(k)] \psi(t-k) dk,$$

wobei $P^*(t) = \int_0^{\infty} \frac{P(k+t)}{P(k)} dF(k)$ und $\psi(t)$ die Lösung des gewöhnlichen Erneuerungsproblems ist.

Aus Satz 13 folgt leicht der

Satz 13a. *Ist bei beliebiger Anfangsalterszusammensetzung der Bestand $H(t)$ beschränkt und gilt entsprechend zu Satz 12 $\psi(t) = \frac{1}{x} + p(t) + \chi(t)$ mit $\lim_{t \rightarrow \infty} \chi(t) = 0$, so ist*

$$\frac{1}{t} \Phi(t) = \frac{1}{x} \cdot \frac{1}{t} \int_0^t H(t) dt + \Phi_1(t) \quad \text{mit} \quad \lim_{t \rightarrow \infty} \Phi_1(t) = 0.$$

Strebt insbesondere $H(t)$ im Mittel gegen einen Wert H_0 , so findet Stabilisierung im Mittel mit dem Werte $\frac{H_0}{x}$ statt.

Beweis. Nach Satz 13 ist

$$(54) \quad \Phi(t) = H(t) - P^*(t) + \int_0^t \frac{H(k)}{x} dk + \int_0^t [H(k) - P^*(k)] p(t-k) dk + \\ + \int_0^t [H(k) - P^*(k)] \chi(t-k) dk - \frac{1}{x} \int_0^t P^*(k) dk.$$

Ist nun in der Tat $0 \leq H(t) \leq M_1$, so ist wegen $|P^*(t)| \leq 1$

$$\left| \int_0^t [H(k) - P^*(k)] p(t-k) dk \right| \leq (M_1 + 1) \int_0^t p(k) dk \leq M_1 + 1.$$

Weiter ist

$$\lim_{t \rightarrow \infty} \left| \frac{1}{t} \int_0^t [H(k) - P^*(k)] \chi(t-k) dk \right| \leq (M_1 + 1) \cdot \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t |\chi(k)| dk = 0.$$

Es wird damit nach (54) unter Beachtung von $\lim_{t \rightarrow \infty} P^*(t) = 0$

$$\lim_{t \rightarrow \infty} \left\{ \frac{\Phi(t)}{t} - \frac{1}{x} \cdot \frac{1}{t} \int_0^t H(k) dk \right\} = 0.$$

Ein Gegenstück zu Satz 13a ist

Satz 13b. *Strebt bei beliebiger Anfangsalterszusammensetzung der beschränkte Gesamtbestand gegen H_0 und gilt $\psi(t) = p(t) + r(t)$ mit beschränktem $r(t)$, so ist*

$$\frac{1}{t} \Phi(t) = \frac{H_0}{t} \int_0^t r(k) dk + \Phi_1(t) \quad \text{mit} \quad \lim_{t \rightarrow \infty} \Phi_1(t) = 0.$$

Es tritt also genau dann Stabilisierung im Mittel ein, wenn $r(t)$ im Mittel konvergiert.

Der Beweis ergibt sich in gleicher Weise wie der von Satz 13a aus Satz 13 durch leichte Abschätzungen unter Beachtung von $\lim_{t \rightarrow \infty} P^*(t) = 0$, der Beziehung (10), sowie der Beschränktheit von $H(t)$, $P(t)$ und $r(t)$.

Unter verschiedenen Einschränkungen könnten wir natürlich noch zu spezielleren Aussagen über $\Phi(t)$ kommen. Wir erwähnen hier lediglich

Satz 13c. *Strebt das beschränkte $H(t)$ gegen H_0 bei konvergentem $\int_0^\infty |H(t) - H_0| dt$, ist $\psi(t) = \frac{1}{x} + p(t) + \chi(t)$ mit $\lim_{t \rightarrow \infty} \chi(t) = 0$ und konvergiert $\int_0^\infty P^*(t) dt$, so ist*

$$\Phi(t) = H_0 \cdot \Psi(t) + c + \Phi_1(t)$$

mit

$$c = H_0 + \frac{1}{x} \int_0^\infty [H(t) - H_0] dt - \frac{1}{x} \int_0^\infty P^*(t) dt, \quad \Psi(t) = \int_0^t \psi(t) dt$$

und

$$\lim_{t \rightarrow \infty} \Phi_1(t) = 0.$$

Der Beweis ergibt sich wieder leicht aus Satz 13 mit Hilfe von (10).

Es ist klar, daß uns für eine Beurteilung des Verlaufs von $\Phi(t)$ die Aussage von Satz 13c nur dann weiterhilft, wenn wir den Verlauf von $\Psi(t)$ beherrschen. Wir werden im nächsten Paragraphen sehen, daß unter gewissen Voraussetzungen $\Psi(t) = \frac{t}{x} + c_1 + \Psi_1(t)$ mit $\lim_{t \rightarrow \infty} \Psi_1(t) = 0$ geschrieben werden kann, wo c_1 eine bekannte Konstante ist. In diesem Falle gibt uns Satz 13c durch $\Phi(t) = H_0 \frac{t}{x} + d + \Phi_1(t)$ mit $\lim_{t \rightarrow \infty} \Phi_1(t) = 0$ bei bekannter Konstanten d ein einfaches Bild vom Verlaufe der Erneuerungszahl.

Wie wir sahen, müssen wir für den Übergang zur Erneuerungsintensität die Totalstetigkeit von $G(t) = H(t) - P^*(t)$ voraussetzen. Bei $G(t) = G(0) + \int_0^t g(t) dt = 0 + \int_0^t g(t) dt$ gewinnen wir dann aus (53) durch eine analog zum Verfahren des § 2 durchzuführende Differentiation

$$(55) \quad \varphi(t) = g(t) + \int_0^t \varphi(k) p(t-k) dk$$

mit der Lösung

$$(56) \quad \varphi(t) = g(t) + \int_0^t g(k) \psi(t-k) dk.$$

Soll nun in der Tat die Erneuerungsintensität einem endlichen Werte zustreben, so können wir aus (55) sofort wörtlich genau so wie beim Beweis zu Satz 5 schließen, daß $g(t)$ oder wenigstens eine äquivalente Funktion zu

$g(t)$ mit wachsendem t gegen Null streben muß. Man überzeugt sich jedoch rasch, daß diese Bedingung durchaus nicht hinreicht.

Sei nun etwa die absolute und quadratische Integrierbarkeit von $g(t)$ verlangt, und sei gemäß Satz 12 $\psi(t)$ darstellbar als $\psi(t) = \frac{1}{x} + p(t) + \chi(t)$ mit $\lim_{t \rightarrow \infty} \chi(t) = 0$; dann liefert (56)

$$\varphi(t) = g(t) + \frac{1}{x} \int_0^t g(k) dk + \int_0^t g(k) p(t-k) dk + \int_0^t g(k) \chi(t-k) dk.$$

Nach (10) streben unter Beibehaltung der Voraussetzung der Quadratintegrierbarkeit von $p(t)$ die beiden letzten Summanden mit wachsendem t gegen Null. Wegen $\lim_{t \rightarrow \infty} P^*(t) = 0$ ist weiter

$$H_0 = \lim_{t \rightarrow \infty} H(t) = \int_0^\infty g(k) dk = \int_0^t g(k) dk + \int_t^\infty g(k) dk,$$

so daß sich ergibt

Satz 14. Ist bei beliebiger Anfangsalterszusammensetzung der Gesamtbestand vorgeschrieben durch $H(t) = P^*(t) + \int_0^t g(k) dk$ mit absolut und quadratisch integrierbarem³⁴⁾ $g(k)$, ist weiter die gewöhnliche Erneuerungsintensität $\psi(t) = \frac{1}{x} + p(t) + \chi(t)$ mit $\lim_{t \rightarrow \infty} \chi(t) = 0$ und quadratintegrierbarem $p(t)$, dann gilt für die zugehörige Erneuerungsintensität

$$\varphi(t) = g(t) + \frac{H_0}{x} + \varphi_1(t) \quad \text{mit} \quad H_0 = \lim_{t \rightarrow \infty} H(t) \quad \text{und} \quad \lim_{t \rightarrow \infty} \varphi_1(t) = 0.$$

Wie wir sehen, macht $\varphi(t)$ schließlich alle Schwankungen von $g(t)$ mit; die Erneuerungsintensität konvergiert genau dann, wenn $\lim_{t \rightarrow \infty} \hat{g}(t) = 0$ für eine äquivalente Funktion $\hat{g}(t)$ zu $g(t)$ ist.

Bei Änderung der Voraussetzungen lassen sich noch weitere derartige Konvergenzsätze bilden. Wir wollen hier jedoch nur noch die Ergebnisse auf den Fall spezialisieren, daß $H(t) \equiv 1$ und die gewöhnliche Erneuerungsintensität konvergent ist. Satz 13a und Satz 13c ergeben dann

Satz 13d. Gilt $\lim_{t \rightarrow \infty} \psi(t) = \frac{1}{x}$, so ist bei beliebiger Anfangsalterszusammensetzung die für die Bestandserhaltung notwendige Zugangsfunktion von der Gestalt $\frac{1}{t} \Phi(t) = \frac{1}{x} + \varphi_1(t)$ mit $\lim_{t \rightarrow \infty} \varphi_1(t) = 0$. Falls $\int_0^\infty P^*(t) dt$ konvergiert, ist

$$\Phi(t) = \Psi(t) + 1 - \frac{1}{x} \int_0^\infty P^*(t) dt + \varphi_2(t) \quad \text{mit} \quad \lim_{t \rightarrow \infty} \varphi_2(t) = 0.$$

³⁴⁾ Falls $\lim_{t \rightarrow \infty} p(t) = 0$ ist, genügt die absolute Integrierbarkeit.

Ist weiter $P^*(t)$ totalstetig, also $P^*(t) = 1 - \int_0^t p^*(t) dt$ mit $p^*(t) \geq 0$, dann ist natürlich $g(t) = p^*(t)$ absolut integrierbar und daher nach Satz 14

Satz 14a. Gilt $\lim_{t \rightarrow \infty} \varphi(t) = \frac{1}{x}$, und $P^*(t) = 1 - \int_0^t p^*(t) dt$, dann ist die für die Bestandserhaltung notwendige Zugangsintensität

$$\varphi(t) = \frac{1}{x} + p^*(t) + \varphi_1(t) \quad \text{mit} \quad \lim_{t \rightarrow \infty} \varphi_1(t) = 0.$$

Das Verhalten von $p^*(t)$ für große t entscheidet also über die Konvergenz von $\varphi(t)$, wobei es natürlich wieder nur auf das Verhalten einer passenden Äquivalenten zu $p^*(t)$ ankommt.

Es bleibt für uns nun noch die Aufgabe, die Integrierbarkeit und Totalstetigkeit von $P^*(t)$ sowie das Grenzverhalten von $p^*(t)$ zu untersuchen.

a) Die Integrierbarkeit von $P^*(t)$.

Ist zunächst das Höchstalter ω endlich, so ist auch $P^*(t) = 0$ für $t > \omega$ und daher $\int_0^\infty P^*(t) dt$ sicher konvergent. Die folgenden Umformungen gelten jedoch durchaus auch für endliches ω .

An Hand von Beispielen können wir uns zunächst rasch überzeugen, daß bei unendlichem Höchstalter $\int_0^\infty P^*(t) dt$ durchaus nicht zu konvergieren braucht³⁵⁾. Um nun einen einfachen Ausdruck für $\int_0^\infty P^*(t) dt$ zu gewinnen, brauchen wir nur (52) direkt zu integrieren. Dies liefert

$$\begin{aligned} \int_0^\infty P^*(t) dt &= \int_{t=0}^\infty \int_{k=0}^\infty \frac{P(t+k)}{P(k)} dF(k) dt = \int_{k=0}^\infty \int_{t=0}^\infty \frac{P(t+k)}{P(k)} dt dF(k) \\ &= \int_{k=0}^\infty \frac{Q(k)}{P(k)} dF(k), \quad Q(k) = \int_k^\infty P(k) dk. \end{aligned}$$

³⁵⁾ Setzen wir etwa $dF(k) = \frac{1}{x} \cdot P(k) dk$ — ein dauernd konstanter Zugang von der Intensität $\frac{1}{x}$ würde einen solchen Altersaufbau hervorrufen —, so ist

$$P^*(t) = \frac{1}{x} \int_0^\infty P(t+k) dk = \frac{1}{x} \int_t^\infty P(k) dk = \frac{1}{x} Q(t),$$

so daß die Konvergenz von $\int_0^\infty P^*(t) dt$ die von $\int_0^\infty t P(t) dt$ voraussetzt.

Die vorgenommene Vertauschung der beiden Integrationen läßt sich unter Beachtung des positiven Charakters aller vorkommenden Funktionen nach Lebesgue³⁶⁾ und Fubini³⁷⁾ direkt rechtfertigen. Wir wollen jedoch für den exakten Beweis der obigen Formel einen kleinen Umweg beschreiten, da derselbe gleichzeitig eine andere bequeme Formel für $P^*(t)$ und sein Integral liefern wird.

Der Gesamtheit gehören $dF(k)$ Mitglieder des Alters k bis $k + dk$ an. Diese Mitglieder können wir uns als aus einer Gesamtheit von $dF_1(k)$ 0-Jährigen der Zeit $t = -k$ bis $t = -k - dk$ entstanden denken, die sich gemäß unserer Absterbeordnung zahlenmäßig vermindert hat; d. h. $dF_1(k) = \frac{dF(k)}{P(k)}$. Wir bilden daher für $k < \infty$

$$(57) \quad F_1(k) = \int_0^k dF_1(k) = \int_0^k \frac{dF(k)}{P(k)}, \quad F_1(\infty) = \lim_{k \rightarrow \infty} F_1(k)$$

und nennen $F_1(k)$ die zum Anfangsbestande gehörige Geburtenfunktion, da $F_1(k)$ die Zahl der 0-Jährigen (= „Geburten“) angibt, aus denen sich die 0- bis k -Jährigen des Anfangsbestandes entwickelt haben.

Wir zeigen nun zunächst die Relation

$$(58) \quad F_1(t) \cdot P(t) \leq 1 \quad \text{mit} \quad \lim_{t \rightarrow \infty} F_1(t) \cdot P(t) = 0.$$

Nach (57) ist nämlich zunächst

$$F_1(t) \cdot P(t) = \int_0^t \frac{P(t)}{P(k)} dF(k) \leq \int_0^t \frac{P(t)}{P(t)} dF(k) \leq 1.$$

Sei nun weiter ein $\varepsilon > 0$ beliebig vorgegeben, dann wählen wir ein $B < \infty$ so groß, daß $\int_B^\infty dF(k) < \varepsilon$ wird. Weiter sei T so groß, daß $\frac{P(T)}{P(B)} < \varepsilon$ ist; insbesondere ist also $T > B$. Es gilt dann für $t > T$

$$\begin{aligned} F_1(t) \cdot P(t) &= \int_0^B \frac{P(t)}{P(k)} dF(k) + \int_B^t \frac{P(t)}{P(k)} dF(k) \leq \int_0^B \frac{P(T)}{P(B)} dF(k) + \\ &\quad + \int_B^t \frac{P(t)}{P(t)} dF(k) < 2\varepsilon. \end{aligned}$$

³⁶⁾ Lebesgue (2), Kap. 11.

³⁷⁾ Ann. 11.

Auf Grund von (58) gewinnen wir aus (52) dann durch partielle Integration

$$P^*(t) = \int_{k=0}^v P(k+t) dF_1(k) = [P(k+t) F_1(k)]_0^v + \int_0^v F_1(k) p(k+t) dk,$$

oder, da wegen $P(k+t) \leq P(k)$ mit (58) auch $\lim_{k \rightarrow \infty} P(k+t) F_1(k) = 0$ ist,

$$P^*(t) = \int_0^v F_1(k) p(k+t) dk.$$

Hieraus ergibt sich nun

$$\int_0^\infty P^*(t) dt = \int_0^\infty \left[\int_0^v F_1(k) p(k+t) dk \right] dt.$$

Der Integrand ist nirgends negativ und liefert zuerst nach t integriert die meß-

bare und nach (58) beschränkte Funktion $\int_0^\infty F_1(k) p(k+t) dt = F_1(k) P(k)$.

Wir können daher nach Fubini unabhängig von der Endlichkeit der mehrfachen Integrale die Integrationsreihenfolge vertauschen und erhalten

$$(59) \quad \int_0^\infty P^*(t) dt = \int_0^v \left[\int_0^\infty F_1(k) p(k+t) dt \right] dk = \int_0^v F_1(k) P(k) dk.$$

Wir zeigen nun weiter in Analogie zu (58):

Bei konvergentem $\int_0^\infty P^*(t) dt$ gilt $\lim_{t \rightarrow \infty} F_1(t) Q(t) = 0$ mit $Q(t) = \int_t^\infty P(k) dk = \int_t^v P(k) dk$. Es ist nämlich wegen des monotonen Nichtfallens von $F_1(k)$

$$0 \leq F_1(t) Q(t) = \int_t^v P(k) F_1(t) dk \leq \int_t^v P(k) F_1(k) dk.$$

Bei konvergentem $\int_0^\infty P^*(t) dt$ ist aber nach (59) $\lim_{t \rightarrow \infty} \int_t^v F_1(k) P(k) dk = 0$ womit die Behauptung klar ist.

Wir wollen nunmehr

$$(60) \quad \int_0^\infty P^*(t) dt = \int_{k=0}^v \frac{Q(k)}{P(k)} dF(k)$$

zeigen.

Gemäß partieller Integration gilt

$$(61) \quad \int_0^t F_1(k) P(k) dk = -F_1(t) Q(t) + \int_0^t \frac{Q(k)}{P(k)} dF(k).$$

Ist nun $\int_0^{\infty} P^*(t) dt$ konvergent, so ergibt sich hieraus bei $t \rightarrow \omega$

$$\int_0^{\infty} P^*(t) dt = \lim_{t \rightarrow \infty} \int_0^t F_1(k) P(k) dk = 0 + \lim_{t \rightarrow \infty} \int_0^t \frac{Q(k)}{P(k)} dF(k).$$

Ist umgekehrt $\int_0^{\infty} \frac{Q(k)}{P(k)} dF(k)$ konvergent, so folgt wegen der Positivität aller vorkommenden Funktionen aus (61) sofort

$$\int_0^t F_1(k) P(k) dk \leq \int_0^t \frac{Q(k)}{P(k)} dF(k) \leq \int_0^{\infty} \frac{Q(k)}{P(k)} dF(k)$$

für jedes t und damit die Konvergenz von $\int_0^{\infty} F_1(k) P(k) dk$, woraus sich aber eben (60) ergab.

Ist schließlich eines der beiden Integrale $\int_0^{\infty} P^*(t) dt$ oder $\int_0^{\infty} \frac{Q(k)}{P(k)} dF(k)$ unendlich, so muß es nach dem geführten Beweis auch das andere sein. so daß (60) in diesem Falle erst recht gilt.

Wir fassen zusammen zu

Satz 15. Setzt man $F_1(k) = \int_0^k \frac{dF(t)}{P(t)}$; so gilt unabhängig von der Endlichkeit der folgenden Integrale:

$$\int_0^{\infty} P^*(t) dt = \int_0^{\infty} \frac{Q(k)}{P(k)} dF(k) = \int_0^{\infty} P(k) F_1(k) dk.$$

(57) läßt sich offenbar nach $F(k)$ auflösen und liefert $F(k) = \int_0^k P(t) dF_1(t)$ und hieraus durch partielle Integration

$$F(k) = F_1(k) P(k) + \int_0^k F_1(t) p(t) dt.$$

Bilden wir hier den Limes $k \rightarrow \omega$, so erhalten wir wegen (58)

$$1 = 0 + \int_0^{\infty} F_1(t) p(t) dt.$$

Die nach Satz 15 für die Konvergenz von $\int_0^\infty P^*(t) dt$ geforderte Endlichkeit von $\int_0^\infty P(k) F_1(k) dk$ ist also höchstens dann eine Verschärfung, wenn $\frac{P(k)}{p(k)} = \frac{1}{\mu(k)}$ für große k nicht beschränkt bleibt. Hieraus folgt aber

Satz 16a. Ist — wie in den meisten praktischen Fällen — $\lim_{x \rightarrow \infty} \mu(x) > 0$, so ist $\int_0^\infty P^*(t) dt$ für jede Anfangsalterszusammensetzung konvergent.

Ein Gegenstück hierzu ist die unmittelbar aus (59) zu ziehende Folgerung

Satz 16b. Ist $F_1(k)$ beschränkt, dann ist $\int_0^\infty P^*(t) dt$ unabhängig von der Absterbeordnung konvergent.

Auch diesem letzteren Falle kommt praktische Bedeutung zu: Beschränktes $F_1(k)$ bedeutet endliches $F_1(\omega)$; $F_1(\omega)$ ist aber die Gesamtzahl der Geburten, aus denen sich die Anfangsgesamtheit entwickelt hat. In der Tat wird man in praktischen Fällen diese Gesamtgeburtenszahl als endlich anzusehen haben.

b) Die Totalstetigkeit von $P^*(t)$.

Versuchen wir zunächst einmal, (52) unter dem Integralzeichen zu differenzieren, so erhalten wir formal

$$\frac{d}{dt} P^*(t) = - \int_0^\infty \frac{p(t+k)}{P(k)} dF(k).$$

Hier müssen wir aber erst noch erklären, welchen Sinn $\int_0^\infty \frac{p(t+k)}{P(k)} dF(k)$ überhaupt haben soll. Ist nämlich $k(F)$ eine Umkehrfunktion zu $F(k)$, so braucht $p(t+k(F))$ bei festem t nicht meßbar in F zu sein³⁸⁾, so daß vielleicht die Lebesguesche Definition³⁹⁾

$$(62) \quad \int_{k=0}^\infty \frac{p(t+k)}{P(k)} dF(k) = \int_{F=0}^1 \frac{p[t+k(F)]}{P[k(F)]} dF$$

nicht durchführbar ist.

Es sei nun tatsächlich $k(F)$ eine Umkehrfunktion zu $F(k)$; d. h. zu jedem F sei ein $k(F)$ so gewählt, daß $F[k(F) - 0] \leq F \leq F[k(F) + 0]$ ist.

³⁸⁾ Carathéodory, § 351, Satz 9.

³⁹⁾ Vgl. Anm. 36, besonders S. 259—261.

$k(F)$ ist eine monoton nichtfallende Funktion mit $k(0) = 0^{40)}$ und $k(1) = \infty$. Wegen $\lim_{k \rightarrow \infty} F(k) = 1$ ist weiter für jedes θ mit $0 < \theta < 1$ $k(\theta)$ eine endliche Größe.

Jedem reellen Werte $x \geq 0$ lassen wir nun alle Punkte der t - F -Ebene mit $t + k(F) = x$ entsprechen. Suchen wir dann für ein vorgegebenes Intervall (x_1, x_2) alle Punkte der t - F -Ebene mit

$$\begin{cases} x_1 \leq t + k(F) \leq x_2, \\ 0 \leq F \leq \theta \text{ bei } 0 < \theta < 1, \end{cases}$$

so ist diese Menge ersichtlich durch Rechteckmengen beliebig genau approximierbar und daher meßbar mit dem Maß $(x_2 - x_1) \theta$. Es entspricht daher überhaupt einer meßbaren x -Menge vom Maße m eine meßbare t - F -Menge vom Maße $m \theta$. Dabei können wir nun auch zu $\theta = 1$ übergehen: Jeder meßbaren x -Menge entspricht eine meßbare t - F -Menge vom gleichen Maß. Da nun $p(x)$ meßbar sein sollte, folgt damit die Meßbarkeit von $p[t + k(F)]$ als Funktion der Variablen t und F . Bei festem $t = t_0$ ist dann nach Fubini⁴¹⁾ $p[t_0 + k(F)]$ für fast alle t_0 ebenfalls meßbar; da auf der anderen Seite die monotone Funktion $P[k(F)]$ sicher meßbar ist, läßt sich das fragliche Stieltjes-Integral tatsächlich durch (62) definieren⁴²⁾.

Um nun zu zeigen, daß (62) auch wirklich die Ableitung von $P^*(t)$ ist, schreiben wir

$$1 - P^*(t) = \int_0^1 \frac{P[k(F)] - P[t + k(F)]}{P[k(F)]} dF = \int_0^1 dF \int_0^t \frac{p[t + k(F)]}{P[k(F)]} dt$$

und hieraus wegen der Meßbarkeit von $p[t + k(F)]$ mit $p \geq 0$ nach Fubini⁴³⁾ und Tonelli⁴⁴⁾

$$1 - P^*(t) = \int_0^t dt \int_0^1 \frac{p[t + k(F)]}{P[k(F)]} dF = \int_0^t dt \int_{k=0}^{\infty} \frac{p(t+k)}{P(k)} dF(k).$$

Es gilt also

Satz 17. Für jede Anfangsalterszusammensetzung ist $P^*(t) = 1 - \int_0^t p^*(t) dt$

mit $p^*(t) = \int_0^{\infty} \frac{p(t+k)}{P(k)} dF(k)$.

⁴⁰⁾ $k(0) = 0$ ist in der Tat stets eine mögliche Definition.

⁴¹⁾ Carathéodory, § 546, Satz 1.

⁴²⁾ In der Ausnahmemenge wird es etwa gleich Null gesetzt.

⁴³⁾ Carathéodory, § 550, Satz 4.

⁴⁴⁾ Carathéodory, § 554, Satz 7.

c) Das Grenzverhalten von $p^*(t)$.

Bei endlichem Höchstalter ist $p^*(t) = 0$ für $t > \omega$; $\lim_{t \rightarrow \infty} p^*(t) = 0$ ist also trivialerweise erfüllt.

Bei unendlichem Höchstalter überzeugt man sich jedoch leicht, daß $p^*(t)$ nicht immer zu konvergieren braucht, auch wenn dies $p(t)$ tut. Wir wollen uns hier mit der Angabe des folgenden Kriteriums begnügen:

Satz 18. *Hinreichend für $\lim_{t \rightarrow \infty} \int_0^t \frac{p(t+k)}{P(k)} dF(k) = 0$ ist die Existenz einer monoton nichtsteigenden gegen Null konvergenten Majoranten $\pi(t)$ für $p(t)$ mit bei passendem $\tau \geq 0$ konvergentem $\int_0^\infty \frac{\pi(k+\tau)}{P(k)} dF(k)$.*

Beweis. Es sei B so groß, daß nach Voraussetzung $\int_B^\infty \frac{\pi(k+\tau)}{P(k)} dF(k) < \varepsilon$ ausfällt, und T so groß, daß $\pi(T) < \varepsilon : F_1(B)$ wird; dann ist für $t \geq \max(T, \tau)$

$$\begin{aligned} \int_0^\infty \frac{p(t+k)}{P(k)} dF(k) &\leq \int_0^B \frac{\pi(t+k)}{P(k)} dF(k) + \int_B^\infty \frac{\pi(t+k)}{P(k)} dF(k) \\ &\leq \pi(T) F_1(B) + \int_B^\infty \frac{\pi(k+\tau)}{P(k)} dF(k) < 2\varepsilon. \end{aligned}$$

Aus Satz 18 folgern wir insbesondere

Satz 18a. *Es ist $\lim_{t \rightarrow \infty} p^*(t) = 0$, wenn eine der folgenden Bedingungen erfüllt ist:*

- a) $\mu(x)$ ist beschränkt,
- b) $F_1(\omega)$ ist endlich und $\lim_{t \rightarrow \infty} p(t) = 0$,
- c) $p(t)$ ist von einem gewissen t an monoton nichtsteigend.

Beweis.

zu a): Mit $\pi(t) = P(t) \cdot \max \mu(t)$ und $\tau = 0$ ist Satz 18 anwendbar;

zu b): für jedes beschränkte $\pi(t)$ ist $\int_0^\infty \frac{\pi(k)}{P(k)} dF(k) = \int_0^\infty \pi(k) dF_1(k)$ konvergent;

zu c): wegen $\int_0^\infty p^*(t) dt = P^*(0) = 1$ ist $p^*(t)$ für fast alle t endlich.

Sei also $p^*(\tau) = \int_0^\infty \frac{p(\tau+k)}{P(k)} dF(k)$ endlich, so erfüllt mit diesem τ jedes $\pi(t)$ die Voraussetzungen des Satzes 18, wenn es von einem gewissen t an mit $p(t)$ identisch ist.

Ein einfaches Beispiel mit $F_1(\infty) = \infty$, wo trotzdem bei jeder Absterbeordnung $\lim_{t \rightarrow \infty} p^*(t) = 0$ gilt, zeigt $F_1(k) = \frac{k}{x}$, das als Geburtenfunktion zu der Altersgliederung $F(k) = \frac{1}{x} \int_0^k P(t) dt$ gehört. In der Tat ist hier nämlich $p^*(t) = \frac{1}{x} \int_0^\infty p(t+k) dk = \frac{1}{x} \cdot P(t)$, so daß $\lim_{t \rightarrow \infty} p^*(t) = \lim_{t \rightarrow \infty} P(t) = 0$ wird.

§ 8.

Nachweis der wahrscheinlichkeitstheoretischen Stabilisierung.

Gemäß den Ausführungen des § 3 findet wahrscheinlichkeitstheoretische Stabilisierung statt, wenn

$$\frac{s^2(t)}{\Phi(t)^2} = \frac{2 \int_0^t \Psi(k) \Psi(t-k) dk + \Psi(t) - \Psi^2(t)}{\Psi^2(t)}$$

für $t \rightarrow \infty$ den Grenzwert 0 annimmt. Dabei ist $\Psi(t) = \int_0^t \psi(t) dt$ mit $\psi(t) = \frac{1}{x} + p(t) + \chi(t)$ bei $\lim_{t \rightarrow \infty} \chi(t) = 0$ zu setzen. Um über $\Psi(t)$ nähere Angaben bezüglich des Verhaltens für große t zu bekommen, die über die mittlere Konvergenz wesentlich hinausgehen, ist es vor allem nötig, das Verhalten von $\int_0^t \chi(t) dt$ zu untersuchen. Dabei werden wir in der gleichen Art vorgehen können wie früher bei der Betrachtung des $r(t)$. Wir wollen daher kurz die Voraussetzungen zusammenfassen, die uns gestatteten, Aussagen über $r(t)$ zu machen:

Es sei $l(x)$ eine stetige Funktion mit der in der unteren Halbebene regulären Transformierten $\lambda(z) = \lambda(\xi - ia) = \lambda(\xi, a) = \lambda_1 - i\lambda_2$, die bei $a \rightarrow +0$ in jedem $\xi \neq 0$ nicht enthaltenden Intervall gleichmäßig einer Grenzfunktion $\lambda(\xi, 0)$ zustrebt. Es sei $\int_{-\infty}^{+\infty} |\lambda(\xi, a)| d\xi$ für $a > 0$ konvergent; $\int_{-1}^{+1} |\lambda(\xi, a)| d\xi$ sei sogar gleichmäßig in a für ein endliches a -Intervall $0 \leq a \leq a_0$.

konvergent. Weiter sei $\lambda_2(\xi, a) \cdot \xi$ für eine gewisse (ξ, a) -Umgebung von $(+0, +0)$ beschränkt. Dann ist

$$(63) \quad \left\{ \begin{array}{l} l(x) = l_1(x) + l_2(x) \quad \text{mit} \\ l_1(x) = \frac{2}{\pi} \int_0^\delta \lambda_2(\xi, 0) \sin \xi x \, d\xi \quad \text{bei beliebigem, festem } \delta \\ \text{und } \lim_{x \rightarrow \infty} l_2(x) = 0. \end{array} \right.$$

Wenden wir diese Schlußweise auf $l(x) = \int_0^x \zeta(t) \, dt$ an. Zunächst ist einmal $\zeta(t) = r(t) - \frac{1}{x}$ und hat damit eine in der unteren Halbebene reguläre Transformierte, da dies für $r(t)$ und jede Konstante zutrifft. Für die stetige Funktion $l(x) = \int_0^x \zeta(t) \, dt$ ergibt sich dann die gleichfalls in der unteren Halbebene reguläre Transformierte

$$\lambda(z) = \frac{a(z) - \frac{1}{x \cdot iz}}{iz}$$

oder gemäß (31)

$$(64) \quad \lambda(z) = \frac{b^2(z)}{iz[1-b(z)]} + \frac{1}{x \cdot z^2}$$

oder mit Hilfe von (44) nach kurzer Rechnung

$$(65) \quad \lambda(z) = -\frac{1+b(z)}{iz} + \frac{1}{x} \cdot \frac{\bar{x} - B(z)}{1-b(z)} \cdot \frac{1}{iz}.$$

Ersichtlich geht $\lambda(\xi, a)$ bei $a \rightarrow +0$ in jedem $\xi = 0$ nicht enthaltenden Intervall tatsächlich gleichmäßig in eine Grenzfunktion über. Aus (64) schließen wir weiter auf Grund der früheren Überlegungen, daß $\int_{-1}^{+\infty} |\lambda(\xi, a)| \, d\xi$ für ein a -Intervall $0 \leq a \leq a_0$ gleichmäßig und daß $\int_{-\infty}^{+\infty} |\lambda(\xi, a)| \, d\xi$ bei $a > 0$ überhaupt konvergent ist. Beim Nachweis der Beschränktheit von $\lambda_2(\xi, a) \cdot \xi$ für eine Umgebung von $\xi = a = +0$ braucht der erste Summand von (65) nicht berücksichtigt zu werden, da sowohl $1+b(z)$ als auch $\frac{\bar{x}}{iz}$ für alle z der unteren Halbebene und der reellen Achse beschränkt sind.

Wir setzen nun wieder $\int_x^\infty P(t) \, dt = Q(x)$, dann hat $Q(x)$ die in der unteren Halbebene sicher reguläre Transformierte

$$C(z) = \frac{B(0) - B(z)}{iz} = \frac{\bar{x} - B(z)}{iz}.$$

Es ist daher wegen (44)

$$(66) \quad \frac{\bar{z} - B(z)}{1 - b(z)} = \frac{iz C(z)}{iz B(z)} = \frac{C(z)}{B(z)}.$$

Es ist daher noch die Beschränktheit von

$$\tau(\xi, a) = -\operatorname{Imag} \left\{ \frac{\xi}{iz} \cdot \frac{C(z)}{B(z)} \right\} = \frac{C_1 B_1 + C_2 B_2}{B_1^2 + B_2^2} \cdot \frac{\xi^2}{\xi^2 + a^2} + \frac{B_1 C_2 - B_2 C_1}{B_1^2 + B_2^2} \cdot \frac{\xi a}{\xi^2 + a^2}$$

zu untersuchen, wobei $C(z) = C_1(\xi, a) - iC_2(\xi, a)$ gesetzt wurde.

Unter den bisher von uns gemachten Annahmen wird nun $\tau(\xi, a)$ im allgemeinen in der Umgebung von $(+0, +0)$ nicht beschränkt sein. Es ist nämlich

$$(67) \quad \tau(\xi, 0) = \frac{C_1 B_1 + C_2 B_2}{B_1^2 + B_2^2} = C_1 \cdot \frac{B_1}{B_1^2 + B_2^2} + C_2 \cdot \frac{B_2}{B_1^2 + B_2^2}.$$

Sei nun etwa $P(t) = \frac{1}{(1+t)^2}$, so ist $Q(t) = \frac{1}{1+t}$ und damit

$$\lim_{a \rightarrow +0} C_1(\xi, a) = C_1(\xi, +0) = \int_0^{\infty} \frac{\cos \xi x}{1+x^2} dx \quad \text{und} \quad C_2(\xi, +0) = \int_0^{\infty} \frac{\sin \xi x}{1+x^2} dx.$$

Wir erkennen $C_1(\xi, 0)$ und $C_2(\xi, 0)$ für diesen Fall als Lommelsche Funktionen, für die die äquivalenten Formeln

$$C_1(\xi, 0) = \int_0^{\infty} \frac{e^{-tx}}{1+x^2} x dx \quad \text{und} \quad C_2(\xi, 0) = \int_0^{\infty} \frac{e^{-tx}}{1+x^2} dx$$

unmittelbar zeigen, daß $\lim_{\xi \rightarrow +0} C_1(\xi, 0) = +\infty$ und $\lim_{\xi \rightarrow +0} C_2(\xi, 0) = \frac{\pi}{2}$ ist. (67) lehrt dann, daß $\tau(\xi, 0)$ in der Umgebung von $\xi = 0$ nicht beschränkt bleibt.

Wir wollen daher an die Absterbeordnung noch die Forderung stellen, daß $\int_0^{\infty} Q(x) dx$ konvergiert. Dann ist $C(\xi, a)$ eine beschränkte Funktion, so daß in der Tat $\xi \cdot \lambda(\xi, a)$ bei $(+0, +0)$ beschränkt bleibt.

Da nun der negative Imaginärteil von $-\frac{1+b(\xi, 0)}{i\xi}$ den Wert $-\frac{1-b_1(\xi, 0)}{\xi}$ hat, erhalten wir schließlich aus (63) mit (65), (66) und (67)

$$l(x) = \int_0^x \chi(t) dt = \frac{2}{\pi} \int_0^{\delta} \left[-1 - b_1 + \frac{1}{x} \cdot \frac{C_1 B_1 + C_2 B_2}{B_1^2 + B_2^2} \right] \cdot \frac{\sin \xi x}{\xi} d\xi + \chi_1(x),$$

wo $\lim_{x \rightarrow \infty} \chi_1(x) = 0$ ist.

Wir schließen hieraus genau so wie oben im § 6 bei der Betrachtung von $r(t)$, daß $l(x)$ im Mittel gegen

$$\lim_{\bar{x} \rightarrow 0} \left[-1 - b_1 + \frac{1}{\bar{x}} \cdot \frac{B_1 C_1 + B_2 C_2}{B_1^2 + B_2^2} \right] = -2 + \frac{C_1(0)}{\bar{x}^2}$$

strebt. Setzen wir $\bar{x}^2 = \int_0^\infty x^2 p(x) dx$, so ist $C_1(0) = \int_0^\infty Q(x) dx = \frac{\bar{x}^3}{2}$. Es strebt also $l(x)$ im Mittel gegen $-2 + \frac{\bar{x}^3}{2\bar{x}^2}$.

Wird zusätzlich die Konvergenz von $\int_0^\infty Q(x) \ln x dx$ vorausgesetzt, so gilt das Grenzwertverhalten nicht nur im Mittel, sondern — wie wir ebenso wie oben bei $r(x)$ zeigen können — sogar exakt.

Unter Beachtung von $\int_0^\infty p(x) dx = 1$ und $\psi(t) = \frac{1}{\bar{x}} + p(t) + \chi(t)$ folgt damit der

Satz 19. Ist $\int_0^\infty x^2 p(x) dx$ konvergent, so strebt $\int_0^x \left[\psi(t) - \frac{1}{\bar{x}} \right] dt$ im Mittel gegen $-1 + \frac{\bar{x}^3}{2\bar{x}^2}$; ist $\int_0^\infty x^2 \ln x p(x) dx$ konvergent, so strebt $\int_0^x \left[\psi(t) - \frac{1}{\bar{x}} \right] dt$ exakt gegen diesen Wert.

Beispiel. Es sei $p(x) = x e^{-x}$; dann ist $\psi(x) = \frac{1}{2} - \frac{1}{2} e^{-2x}$, $\bar{x} = 2$, $\bar{x}^2 = 6$. Also ist $\psi(x) - \frac{1}{\bar{x}} = -\frac{1}{2} e^{-2x}$ und daraus $\int_0^\infty \left[\psi(x) - \frac{1}{\bar{x}} \right] dx = -\frac{1}{4}$. In der Tat ist $-1 + \frac{\bar{x}^3}{2\bar{x}^2} = -1 + \frac{6}{8} = -\frac{1}{4}$.

Nehmen wir nun in der Tat an, daß $\int_0^\infty x^2 p(x) dx$ konvergiert, dann gilt zunächst nach Satz 12

$$\psi(t) = \frac{1}{\bar{x}} + p(t) + \chi(t) \quad \text{mit} \quad \lim_{t \rightarrow \infty} \chi(t) = 0$$

und nach Satz 19

$$\Psi(t) = \int_0^t \psi(k) dk = \frac{t}{\bar{x}} + c + \gamma(t) \quad \text{mit} \quad c = -1 + \frac{\bar{x}^3}{2\bar{x}^2}$$

$$\text{und} \quad \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t \gamma(k) dk = 0.$$

Bilden wir jetzt die Faltung $\int_0^t \Psi(k) \varphi(t-k) dk$, so treten dabei die folgenden Glieder auf:

$$h_1(t) = \int_0^t k p(t-k) dk = Q(t) + t - \bar{x},$$

$$h_2(t) = \int_0^t k \chi(t-k) dk = (c-1)t + \int_0^t \gamma(k) dk + \bar{x} - Q(t),$$

$$h_3(t) = \int_0^t \gamma(k) p(t-k) dk = \gamma(t) - \int_0^t P(t-k) \chi(k) dk + \\ - \int_0^t P(t-k) p(k) dk + P(t) \cdot c,$$

$$h_4(t) = \int_0^t \gamma(k) \chi(t-k) dk.$$

Ersetzen wir eine Funktion, die mit wachsendem t gegen Null strebt, durch das Symbol $o(1)$, so ist unter Beachtung von (10) ersichtlich

$$h_1(t) = t - \bar{x} + o(1),$$

$$\frac{1}{t} h_2(t) = c - 1 + o(1),$$

$$h_3(t) = \gamma(t) + o(1),$$

$$\frac{1}{t} \gamma(t) = \frac{1}{t} \int_0^t \chi(t) dt + \frac{1-c-P(t)}{t} = o(1),$$

so daß wir für $s^2(t)$ gemäß (23) nach kurzer Rechnung erhalten

$$(68) \quad s^2(t) = \frac{t}{\bar{x}} \cdot \left[2c + 1 - 2\gamma(t) - \bar{x} \cdot \frac{\gamma^2(t)}{t} + 2\bar{x} \frac{h_4(t)}{t} + o(1) \right].$$

Andererseits ist

$$(69) \quad \Psi(t) = \frac{t}{\bar{x}} [1 + o(1)],$$

so daß wir erhalten

$$(70) \quad \frac{s^2(t)}{\Psi^2(t)} = \bar{x} \left[\frac{2c+1}{t} - 2 \frac{\gamma(t)}{t} - \bar{x} \frac{\gamma^2(t)}{t^2} + 2\bar{x} \frac{h_4(t)}{t^2} + o(1) \right].$$

Nun ist

$$\left| \frac{h_4(t)}{t^2} \right| = \left| \frac{1}{t} \int_0^t \frac{\gamma(k)}{t} \chi(t-k) dk \right| \leq \frac{1}{t} \int_0^t \left| \frac{\gamma(k)}{k} \right| \cdot |\chi(t-k)| dk$$

und damit wegen $\frac{\gamma(k)}{k} = o(1)$ und $\chi(k) = o(1)$ auch $\frac{h_4(t)}{t^2} = o(1)$. Es ist also nach (70)

$$\frac{s^2(t)}{\Psi^2(t)} = o(1),$$

womit die wahrscheinlichkeitstheoretische Stabilisierung nachgewiesen ist.

Wenn $\gamma(t)$ sogar exakt gegen Null geht wie im Falle der Konvergenz von $\int_0^\infty x^2 \ln x p(x) dx$, so ist ersichtlich sogar $\frac{1}{t} h_4(t) = o(1)$, und wir erhalten aus (68) und (69)

$$\frac{s^2(t)}{\Psi(t)} = 2c + 1 + o(1) = \frac{\bar{x}^2 - \bar{x}^2}{\bar{x}^2} + o(1).$$

Wir fassen zusammen zu

Satz 20. Ist $\int_0^\infty x^2 p(x) dx$ konvergent, so ist wahrscheinlichkeitstheoretische Stabilisierung vorhanden.

Ist $\int_0^\infty x^2 \ln x p(x) dx$ konvergent, so ist $\frac{s^2(t)}{\Psi(t)} = \frac{\bar{x}^2 - \bar{x}^2}{\bar{x}^2} + o(1)$.

Es geht dann also der Quotient aus Streuung und Erwartungswert gegen Null wie $\frac{C}{\sqrt{t}}$ bei $C = \sqrt{\frac{\bar{x}^2 - \bar{x}^2}{\bar{x}^2}}$.

Literaturverzeichnis.

I.

- S. Bochner: Vorlesungen über Fouriersche Integrale. Leipzig 1932.
 C. Carathéodory: Vorlesungen über reelle Funktionen. Leipzig und Berlin 1927.
 G. Doetsch: Theorie und Anwendung der Laplace-Transformation. Berlin 1937.
 H. Hadwiger:
 [1] Zur Berechnung der Erneuerungsfunktion nach einer Formel von V. A. Kostitzin. MVSV (Mitteilungen der Vereinigung schweizerischer Versicherungsmathematiker), Heft 34 (1937).
 [2] Zur Frage des Beharrungszustandes bei kontinuierlich sich erneuernden Gesamtheiten. A. f. m. W. S. (Archiv für mathematische Wirtschafts- und Sozialforschung), Band V, Heft 1 (1939).
 L. Herbelot: Application d'un théorème d'analyse à l'étude de la réparation par âge dans les milieux animés dont l'effectif est constant. Bulletin trim. de l'Inst. d. Actuaires Français, 1909.
 K. Kohler: Jährliche Ausfallmenge eines Bestandes von Holzmasten. A. f. m. W. S., Band II, Heft 1 (1936).
 V. A. Kostitzin: Applications des équations intégrales (applications statistiques). Mémoires d. sciences mathématiques, Heft LXIX. Paris 1935.
 H. Lebesgue:
 [1] Leçons sur les séries trigonométriques. Paris 1906.
 [2] Leçons sur l'intégration. Paris 1928.
 Chr. Moser:
 [1] Integralgleichungen und sich erneuernde Gesamtheiten. Ber. d. 9. intern. Kongresses f. Versicherungswissenschaft, Stockholm 1930, Band 3.
 [2] Beiträge zur Darstellung von Vorgängen und des Beharrungszustandes bei einer sich erneuernden Gesamtheit. MVSV, Heft 21 (1926).
 R. Paley und N. Wiener: Fourier transforms in the complex domain. Amer. Math. Soc. Coll. Public., Band XIX (1934).

H. Richter:

- [1] Die Konvergenz der Erneuerungsfunktion. Blätter für Vers. mathem. u. verwandte Gebiete, Band 5, Heft 1 (1940).
- [2] Eine Bemerkung zum Erneuerungsproblem. A. f. m. W. S. Band VI, Heft 3 (1940).

R. Rissler: Sur une application d'une équation fonctionnelle à un problème d'assurance. Bull. trim. de l'Inst. des Actuaires Français, 1912.

S. Saks: Théorie de l'intégrale. Warschau 1933.

H. Schulthess: Über das Erneuerungsproblem bei Verwendung eines analytischen Sterbegesetzes. MVSU, Heft 33 (1937).

E. Zwinggi:

- [1] Das Problem der Erneuerung. Festgabe Moser, Bern 1931.
- [2] Bemerkungen zum Erneuerungsproblem. MVSU, Heft 36 (1938).

II.

H. Hadwiger: Untersuchungen über das asymptotische Verhalten rekurrenter Zahlenreihen. MVSU, Heft 35 (1938).

H. Kreis: Stabilität einer sich jährlich erneuernden Gesamtheit. MVSU, Heft 32 (1936).

A. Maret: Untersuchungen über diskontinuierlich sich erneuernde Gesamtheiten. Diss. Bern 1936.

W. Moeschler: Abbau und Erneuerung des Bestandes einer Sterbekasse. Festgabe Moser, Bern 1931.

H. Münzner:

- [1] Die Erneuerung von Gesamtheiten. A. f. m. W. S., Band IV, Heft 1 (1938).
- [2] Der Grenzwert der Erneuerungszahlen. A. f. m. W. S., Band V, Heft 1 (1939).

H. Münzner und H. Schwarz: Ein Zusammenhang zwischen Erneuerungszahlen und dem Moivreschen Problem. A. f. m. W. S., Band 6, Heft 1 (1940).

W. Saxer: Zur Frage des Beharrungszustandes. MVSU, Heft 27 (1932).

E. Zwinggi: Zum Problem der Erneuerung. Blätter für Vers. math. u. verw. Gebiete, Band 2, Heft 1 (1931).

III.

W. Friedli: Über die Stabilität der gegenseitigen Hilfskassen. Zeitschr. f. schweiz. Statistik u. Volkswirtschaft, 1927.

W. Moeschler: Untersuchungen über Eintrittsgewinn und Fehlbetrag einer Versicherungskasse. MVSU, Heft 30 (1935).

G. Schärflin: Über die Höhe der finanziellen Belastung, welche durch die Altersversorgung der eidgenössischen Beamten und Angestellten voraussichtlich hervorgerufen wird. Zeitschr. f. schweiz. Statistik u. Volkswirtschaft, 1899.

W. Thalman: Der Beharrungszustand in der sozialen Unfallversicherung. Festgabe Moser, Bern 1931.

H. Wyma:

- [1] Die Bemessung der Reserven schweizerischer Krankenkassen. Zeitschr. f. schweiz. Statistik u. Volkswirtschaft, 1927.
- [2] Lage, Entwicklung und Beharrungszustand der eidgenössischen Versicherungskasse. MVSU, Heft 24 (1929).

(Eingegangen am 29. 10. 1940.)

Zur Darstellungstheorie der Raumgruppen*).

Von

Georg Wintgen in Leipzig.

Eine Raumgruppe ist eine anendliche diskrete Gruppe von euklidischen Bewegungen und Umlegungen des dreidimensionalen Raumes, welche keinen Punkt und keine Gerade oder Ebene invariant läßt. Es gibt 230 verschiedene, d. h. nicht affin ineinander transformierbare Raumgruppen. Die in einer Raumgruppe \mathfrak{G} enthaltenen Translationen bilden einen Abelschen Normalteiler Γ , der von drei linear unabhängigen Translationen erzeugt wird, also die Struktur \mathfrak{C}_{000} hat. Die Faktorgruppe $\mathfrak{H} = \mathfrak{G}/\Gamma$ ist jeweils eine von 18 verschiedenen endlichen Gruppen. \mathfrak{G} ist eine Erweiterung von \mathfrak{C}_{000} mit \mathfrak{H} , man kann also die Elemente von \mathfrak{H} als Automorphismen von \mathfrak{C}_{000} auffassen. So kommt man zu einer ganzzahligen unimodularen Darstellung von \mathfrak{H} durch dreireihige Matrizes. Diese Darstellung definiert die zu \mathfrak{G} gehörige arithmetische Kristallklasse [2] [6]**). Es gibt 73 arithmetische Kristallklassen, wenn man zwei Darstellungen, die sich ganzzahlig unimodular ineinander transformieren lassen, als nicht verschieden ansieht. Wendet man Γ auf einen festen Punkt 0 des Raumes an, so erhält man ein dreidimensionales Gitter. Zerlegt man jede Bewegung oder Umlegung aus \mathfrak{G} in eine Translation und in eine Drehung oder Umlegung mit dem festen Punkt 0, so bilden die rotativen Bestandteile für sich eine Gruppe, die zu \mathfrak{G} gehörige Punktgruppe, die zu \mathfrak{H} isomorph ist. Wählt man die Gittervektoren als Koordinatenvektoren, so werden die Elemente der Punktgruppe unimodulare ganzzahlige lineare Vektortransformationen, und man kommt so zu einer Darstellung von \mathfrak{H} durch dreireihige Matrizes, die mit der obigen genau übereinstimmt (vgl. v. d. Waerden [1]).

Im folgenden sollen nun alle beschränkten irreduziblen Darstellungen der 230 Raumgruppen aufgestellt werden, aus denen sich nach einem Satz von J. von Neumann alle übrigen beschränkten Darstellungen zusammensetzen.

In § 1 werden wir zeigen, daß sich bei passender Wahl der Basisvektoren im Darstellungsraum alle beschränkten Darstellungen von \mathfrak{G} auf eine ganz bestimmte, im wesentlichen nur von der arithmetischen Kristallklasse abhängige Form reduzieren. Dann zeigen wir in § 2, daß jedes System von Matrizes, das man in dieser Form ansetzt, auch wirklich eine Darstellung

*) D 15. Diese Arbeit wurde von der philosophischen Fakultät der Universität Leipzig als Dissertation angenommen.

**) Fußnoten beziehen sich auf das Literaturverzeichnis am Ende der Arbeit.

von \mathfrak{G} liefert. In § 3 suchen wir unter den so erhaltenen Darstellungen die irreduziblen und die untereinander inäquivalenten aus. In § 4 geben wir eine gruppentheoretische, in § 5 eine geometrische Übersicht über diese Darstellungen, in § 6 endlich eine Zusammenstellung der arithmetischen Kristallklassen und ein Beispiel.

In der Basiswahl folgen wir F. Seitz [3], jedoch bedienen wir uns eines Kunstgriffes, durch den die Darstellungen zweier Raumgruppen derselben arithmetischen Kristallklasse auf dieselbe Form gebracht werden. Sei, um dieses zu erläutern, \mathfrak{R} der Raum, in dem \mathfrak{G} eine beschränkte Darstellung erfährt, dann zerfällt \mathfrak{R} unter Γ in eindimensionale Teilräume $\{u\}$ derart, daß für $t \in \Gamma$

$$tu = e^{2\pi i l(t)} u = \lambda(t)u$$

ist, wobei $l(t)$ eine von u abhängige Linearform der Komponenten des Vektors ist. Wir greifen einen der Vektoren u heraus und wählen nun die h Vektoren $\alpha_i u$ als Basis, wobei die α_i die h Elemente der Punktgruppe sind. Die $\{\alpha_i u\}$ sind auch unter Γ invariant, wobei

$$t(\alpha_i u) = \lambda_i(t) (\alpha_i u) = e^{2\pi i l_i(t)} (\alpha_i u)$$

ist. Diese Wahl der Basisvektoren ist nun zunächst nur für eine solche Gruppe \mathfrak{G} möglich, die die Elemente der Punktgruppe unter ihren Elementen enthält, also für die sogenannten symmorphen Raumgruppen der arithmetischen Kristallklassen [6]. Für die anderen Raumgruppen erreichen wir aber dasselbe durch den erwähnten Kunstgriff. Sei a, α_i die euklidische Transformation aus \mathfrak{G} , die entsteht, wenn man zunächst die Drehung oder Umlegung α_i mit dem Fixpunkt 0 ausführt und sodann die Translation a , und sei im besonderen dieses Element a, α_i ein Repräsentant der Restklasse nach Γ , die dem Element α_i der Punktgruppe entspricht, so betrachten wir zunächst den Vektor $(a, \alpha_i)u$. Er spannt einen unter Γ invarianten Teilraum von \mathfrak{R} aus und es sei

$$t(a, \alpha_i u) = \lambda_i(t) (a, \alpha_i u).$$

Als r -ten Basisvektor wählen wir dann den Vektor

$$\lambda_i(a_i)^{-1} (a, \alpha_i)u.$$

Damit haben wir de facto dieselbe Basiswahl wie oben durchgeführt.

Sei α ein Element der Punktgruppe und t eine Translation, so liefert α auf t angewandt eine neue Translation t' , die wir mit

$$t' = \alpha t \alpha^{-1}$$

bezeichnen. Für zwei Elemente $t\alpha$ und $s\beta$ aus \mathfrak{G} gilt das Multiplikationsgesetz:

$$t\alpha s\beta = t\alpha s\alpha^{-1}\alpha\beta = t's'\alpha\beta.$$

Unter Darstellung schlechthin verstehen wir im folgenden immer eine beschränkte Darstellung.

§ 1.

Sei $\theta(t\alpha_r)$ eine Darstellung von \mathfrak{G} in einem Darstellungsraum \mathfrak{R} , $\{u\}$ ein unter Γ invarianter Teilraum von \mathfrak{R} , t eine Translation aus Γ und sei

$$(1) \quad tu = \lambda(t)u = e^{2\pi i t(t)}u,$$

$$(2) \quad l(t) = n_1 \varphi_1 + n_2 \varphi_2 + n_3 \varphi_3,$$

wobei die φ_r modulo 1 bestimmte reelle Zahlen und die n_r die Gitterkoordinaten von t seien. Sei ferner h die Ordnung von \mathfrak{G} und $a_r \alpha_r$ ein Repräsentant der r -ten Restklasse nach Γ ($a_1 \alpha_1 = \varepsilon$ = Einselement von \mathfrak{G}), so ist

$$ta_r \alpha_r = a_r t \alpha_r = a_r \alpha_r t^{a_r^{-1}}$$

und folglich

$$(3) \quad t[(a_r \alpha_r)u] = \lambda(t^{a_r^{-1}})(a_r \alpha_r)u.$$

Demgemäß bilden wir die h Vektoren

$$(4) \quad u_r = u_{a_r} = \lambda_r^{-1}(a_r)(a_r \alpha_r)u,$$

wobei

$$(5) \quad \lambda_r(t) = \lambda(t^{a_r^{-1}}).$$

Satz 1. Diese h Vektoren u_r ($r = 1, 2, \dots, h$) spannen einen n -dimensionalen Teilraum \mathfrak{S} von \mathfrak{R} auf ($n \leq h$), der unter \mathfrak{G} invariant ist.

Beweis. $ta_r u_r = ta_r \lambda_r^{-1}(a_r) a_r \alpha_r u = \lambda' s a_\mu \alpha_\mu u$, wobei λ' eine komplexe Zahl, s eine Translation aus Γ und $a_\mu \alpha_\mu$ ein Restklassenrepräsentant ist. Ziehen wir nun noch aus λ' den Faktor $\lambda_\mu^{-1}(a_\mu)$ heraus, so erhalten wir

$$ta_r u_r = \lambda'' s \lambda_\mu^{-1}(a_\mu) a_\mu \alpha_\mu u = \lambda'' s u_\mu = \lambda''' u_\mu \in \mathfrak{S}.$$

Es kann $n < h$ sein, d. h. die h Vektoren u_r brauchen nicht linear unabhängig zu sein. Ferner sind die h Funktionen $\lambda_r(t)$ nicht immer alle voneinander verschieden. Es gilt aber

Satz 2. Seien $\lambda_1(t), \lambda_2(t), \dots, \lambda_d(t)$ die d verschiedenen unter den Funktionen $\lambda_r(t)$, dann sind die d Vektoren u_1, \dots, u_d linear unabhängig.

Beweis. Gesetzt, schon c linear unabhängige unter ihnen würden genügen, um die übrigen auszudrücken, so hätte man für $r > c$

$$u_r = \sum_{\mu=1}^c \kappa_{r\mu} u_\mu.$$

Wenden wir nun links und rechts $t \in \Gamma$ an, so erhalten wir:

$$\lambda_r(t) u_r = \sum_{\mu=1}^c \kappa_{r\mu} \lambda_\mu(t) u_\mu,$$

$$\sum_{\mu=1}^c \kappa_{r\mu} \lambda_r(t) u_\mu = \sum_{\mu=1}^c \kappa_{r\mu} \lambda_\mu(t) u_\mu,$$

was der eindeutigen Darstellung von u , als Linearkombination der linear unabhängigen u_μ ($\mu = 1, 2 \dots c$) widerspricht.

Wir konstruieren jetzt eine Basis für den Raum \mathfrak{S} , die auch für den Fall $n < h$ brauchbar ist. Diejenigen unter den h Vektoren u_μ , die zum Eigenwert $\lambda_1 = \lambda$ der Translationsgruppe gehören, spannen einen Teilraum A_1 von \mathfrak{S} auf; s unter ihnen, u_{11}, \dots, u_{1s} , seien eine Basis von A_1 , dann bilden wir zu jedem der d verschiedenen λ_μ die s Vektoren

$$(6) \quad u_{\mu\varrho} = \lambda_\mu (a_\mu)^{-1} (a_\mu \alpha_\mu) u_{1\varrho} \quad (\varrho = 1, 2, \dots, s).$$

Satz 3. Die s Vektoren $u_{\mu\varrho}$ ($\varrho = 1, 2, \dots, s, \mu$ fest) sind linear unabhängig, sie spannen deshalb einen s -dimensionalen Teilraum A_μ von \mathfrak{S} auf. A_μ gehört zu λ_μ , d. h. er besteht aus lauter Eigenvektoren zum Eigenwert $\lambda_\mu(t)$.

Beweis. Aus $\sum_{\varrho} \kappa_{\varrho} u_{\mu\varrho} = 0$ folgt

$$\sum_{\varrho} \kappa_{\varrho} \lambda_\mu^{-1} (a_\mu) a_\mu \alpha_\mu u_{1\varrho} = (a_\mu \alpha_\mu) \lambda_\mu^{-1} (a_\mu) \sum_{\varrho} \kappa_{\varrho} u_{1\varrho} = 0$$

oder

$$\sum_{\varrho} \kappa_{\varrho} u_{1\varrho} = 0,$$

da $a_\mu \alpha_\mu$ eine nicht singuläre Transformation der Vektoren von \mathfrak{S} bedeutet und da $\lambda_\mu^{-1} (a_\mu) \neq 0$. Aus der linearen Unabhängigkeit der $u_{1\varrho}$ folgt also die lineare Unabhängigkeit der $u_{\mu\varrho}$.

Um zu zeigen, daß $A_\mu = (u_{\mu 1}, \dots, u_{\mu s})$ zum Eigenwert $\lambda_\mu(t)$ gehört, zeigen wir es für die Basisvektoren:

$$\begin{aligned} t u_{\mu\varrho} &= t \lambda_\mu^{-1} (a_\mu) (a_\mu \alpha_\mu) u_{1\varrho} = \lambda_\mu^{-1} (a_\mu) a_\mu \alpha_\mu t^{a_\mu^{-1}} u_{1\varrho} \\ &= \lambda_\mu^{-1} (a_\mu) a_\mu \alpha_\mu \lambda (t^{a_\mu^{-1}}) u_{1\varrho} = \lambda_\mu(t) u_{\mu\varrho}. \end{aligned}$$

Satz 4. Es ist $\mathfrak{S} = A_1 + A_2 + \dots + A_d$. \mathfrak{S} hat also die Dimension $n = s \cdot d$, und die Vektoren $u_{\mu\varrho}$ ($\mu = 1, 2, \dots, d, \varrho = 1, 2, \dots, s$) bilden eine Basis von \mathfrak{S} .

Beweis. Wir zeigen, daß jeder der h Vektoren u , die \mathfrak{S} aufspannen, in einem der Räume A_μ enthalten ist. Angenommen u gehöre zu λ_μ , dann gehört

$$v = (a_\mu \alpha_\mu)^{-1} u,$$

wegen

$$t v = t (a_\mu \alpha_\mu)^{-1} u = (a_\mu \alpha_\mu)^{-1} t^{a_\mu} u = (a_\mu \alpha_\mu)^{-1} \lambda \left[(t^{a_\mu})^{a_\mu^{-1}} \right] u = \lambda(t) v$$

zu $\lambda_1 = \lambda$. Aber v läßt sich in der Gestalt

$$v = \lambda' s a_s \alpha_s u$$

schreiben, wobei λ' eine komplexe Zahl, s eine Translation aus Γ und $a_s \alpha_s$ ein Restklassenrepräsentant ist. Da v zu λ_1 gehört, so gehört auch

$$\lambda (a_s^{-1}) a_s \alpha_s u = u,$$

zu λ_1 , woraus sich ergibt, daß dieser Vektor und somit auch v selbst zu A_1 gehört; das bedeutet aber, daß v eine Linearkombination der u_{1q} ist:

$$v = (a_\mu \alpha_\mu)^{-1} u_\mu = \sum_q \kappa_q u_{1q},$$

also:

$$u_\mu = \sum_q \kappa_q (a_\mu \alpha_\mu) u_{1q} = \sum_q \kappa'_q u_{\mu q} \in A_\mu, \text{ q. e. d.}$$

Damit ist eine Zerlegung von \mathfrak{R} in Räume \mathfrak{S} der Dimension $n \leq h$ durchgeführt. Wir studieren nun die Darstellung, die \mathfrak{G} in \mathfrak{S} erfährt. Es gilt:

Satz 5. Die Räume A_μ werden durch die Elemente $a\alpha$ von \mathfrak{G} lediglich permutiert, wobei die Permutation nur von α , also nur von der Restklasse nach Γ abhängt.

Beweis. Wir zeigen zunächst, daß in \mathfrak{S} jeder Eigenvektor einer Translation aus \mathfrak{G} in einem der Räume A_μ enthalten ist. Sei

$$tv = \lambda v,$$

so zerlegen wir v in Komponenten nach den einzelnen Räumen und es ergibt sich links:

$$tv = t \sum_{\mu=1}^d \kappa_\mu v_\mu = \sum_{\mu=1}^d \kappa_\mu tv_\mu = \sum_{\mu=1}^d \kappa_\mu \lambda_\mu(t) v_\mu,$$

rechts:

$$\lambda v = \lambda \sum_{\mu=1}^d \kappa_\mu v_\mu = \sum_{\mu=1}^d \kappa_\mu \lambda v_\mu.$$

Wegen der Verschiedenheit der λ_μ können diese Summen nur dann gleich sein, wenn sie sich auf einen Summanden reduzieren, wenn also

$$\lambda v = \kappa_\mu \lambda v_\mu$$

gilt oder

$$v = \kappa_\mu v_\mu \in A_\mu.$$

Ist nun v ein Vektor aus A_μ , also ein Eigenvektor der Translationsgruppe zum Eigenwert $\lambda_\mu(t)$, so ist auch $a\alpha v$ ein Eigenvektor, der wegen

$$t(a\alpha v) = a\alpha t^{a^{-1}} v = a\alpha \lambda_\mu(t^{a^{-1}}) v = \lambda_\mu t^{a^{-1}} a\alpha v$$

zum nur von μ und α abhängigen Eigenwert

$$(7) \quad \lambda_\mu(t^{a^{-1}}) = \lambda_{P_\alpha(\mu)}(t) = \lambda_\nu(t)$$

gehört. Nach dem obigen ist dieser einer der d verschiedenen Eigenwerte λ_μ , und man erkennt, daß alle Vektoren v aus A_μ durch $a\alpha$ in denselben nur von μ und α abhängigen Raum $A_{P_\alpha(\mu)}$ transformiert werden. $\mu \rightarrow \nu = P_\alpha(\mu)$

definiert die α zugeordnete Permutation der Räume A_r . Satz 5 ist hiermit bewiesen.

Wir können nun ansetzen:

$$(8) \quad t\alpha u_{\mu\nu} = \lambda_r(t) \sum_{\tau} u_{r\tau} D_{\tau\nu}^{\mu}(\alpha) \quad (\nu = P_{\alpha}(\mu)).$$

Damit haben wir s -reihige Matrizes $D^{\mu}(\alpha) = (D_{\tau\nu}^{\mu}(\alpha))$ definiert, die nur von α abhängen; denn für $t' \subset \Gamma$ gilt:

$$t'(t\alpha u_{\mu\nu}) = \lambda_r(t) \sum_{\tau} \lambda_{r'}(t') u_{r\tau} D_{\tau\nu}^{\mu}(\alpha) = \lambda_{r'}(t') \sum_{\tau} u_{r\tau} D_{\tau\nu}^{\mu}(\alpha)$$

und die Translation $t't$ (die Summe von t' und t !) durchläuft alle Translationen von Γ . Sind umgekehrt die $D^{\mu}(\alpha)$, die Permutationen $P_{\alpha}(\mu)$ und die $\lambda_r(t)$ bekannt, so ist durch (8) die Darstellung von \mathfrak{G} bestimmt. Sei z. B. die Anzahl d der Räume A_r gleich drei und sei $P_{\alpha} = (123)$, so wird das Element $t\alpha$ gemäß (8) durch die Matrix

$$(9) \quad t\alpha \rightarrow \begin{pmatrix} 0 & 0 & \boxed{\lambda_1(t) D^3(\alpha)} \\ \boxed{\lambda_2(t) D^1(\alpha)} & 0 & 0 \\ 0 & \boxed{\lambda_3(t) D^2(\alpha)} & 0 \end{pmatrix}$$

dargestellt. Wir beachten, daß die λ_r den Zeilenindex r , die D^{μ} den Spaltenindex μ tragen.

Wir wollen jetzt die $P_{\alpha}(\mu)$ und die $D^{\mu}(\alpha)$ genauer betrachten. Zu diesem Zwecke können wir von den Translationen absehen und unser Augenmerk auf die zu der Darstellung (8) von \mathfrak{G} gehörige Darstellung von \mathfrak{H} richten, die durch

$$(10) \quad \alpha u_{\mu\nu} = \sum_{\tau} u_{r\tau} D_{\tau\nu}^{\mu}(\alpha) \quad (P_{\alpha}(\mu) = \nu)$$

definiert ist. Wir müssen noch beweisen, daß dies wirklich eine Darstellung von \mathfrak{H} ist. Dies ist dann und nur dann der Fall, wenn dem Produkt zweier Elemente β und α aus \mathfrak{H} mit (10) und

$$(11) \quad \beta u_{rs} = \sum_{\sigma} u_{\sigma\sigma} D_{\sigma r}^s(\beta) \quad (P_{\beta}(r) = s)$$

das Produkt der zugeordneten linearen Transformationen entspricht, wenn also gilt:

$$(12) \quad (\beta\alpha) u_{\mu\nu} = \sum_{\sigma} \sum_{\tau} u_{\sigma\sigma} D_{\sigma\tau}^r(\beta) D_{\tau\nu}^{\mu}(\alpha).$$

Dies sieht man aber nach kurzer Rechnung ein: Sei (11) entsprechend

$$s\beta u_{rs} = \lambda_x(s) \sum_{\sigma} u_{\sigma\sigma} D_{\sigma r}^s(\beta) \quad (P_{\beta}(r) = s),$$

so ist

$$s\beta t\alpha u_{\mu\nu} = s\beta \lambda_r(t) \sum_{\tau} u_{r\tau} D_{\tau\nu}^{\mu}(\alpha) = \lambda_x(s) \lambda_r(t) \sum_{\sigma} \sum_{\tau} u_{\sigma\sigma} D_{\sigma r}^s(\beta) D_{\tau\nu}^{\mu}(\alpha).$$

Nun ist aber

$$s\beta t\alpha = st^j\beta\alpha$$

und, wie aus zweimaliger Anwendung von (7) folgt,

$$\lambda_x(s)\lambda_r(t) = \lambda_r(st^{j-1}t) = \lambda_r((st^j)^{j-1}) = \lambda_x(st^j).$$

Also haben wir

$$(13) \quad st^j\beta\alpha u_{\mu\varrho} = \lambda_x(st^j) \sum_{\alpha} \sum_{\tau} u_{x\alpha} D'_{\alpha\tau}(\beta) D''_{\tau\alpha}(\alpha).$$

Sieht man jetzt von der Translation ab, so kommt man genau zu (12), q. e. d.

Wir betrachten also jetzt diese Darstellung (10) von \mathfrak{S} . Wir ersetzen in (10) α durch $\beta\alpha$ und erhalten:

$$(\beta\alpha)u_{\mu\varrho} = \sum_{\alpha} u_{\lambda\alpha} D''_{\alpha\varrho}(\beta\alpha) \quad (P_{\beta\alpha}(\mu) = \lambda).$$

Der Vergleich mit (12) ergibt:

$$(14) \quad P_{\beta\alpha}(\mu) = P_{\beta}(P_{\alpha}(\mu))$$

$$(15) \quad D''(\beta\alpha) = D^{P_{\alpha}(\mu)}(\beta) D''(\alpha).$$

(14) bedeutet, daß die Zuordnung $\alpha \rightarrow P_{\alpha}(\mu)$ eine Darstellung von \mathfrak{S} durch eine Permutationsgruppe ist, (15) wird dazu dienen, die Matrizes $D''(\alpha)$ zu berechnen.

Diejenigen Elemente β von \mathfrak{S} , die A_1 invariant lassen, für die also $P_{\beta}(1) = 1$ gilt, bilden eine Untergruppe \mathfrak{R} von \mathfrak{S} vom Index d und die Matrizes $D^{\beta}(\beta) = D(\beta)$ bilden eine Darstellung von \mathfrak{R} in A_1 .

Wir werden nun zeigen, daß durch Angabe von \mathfrak{R} und $D(\beta)$ für $\beta \in \mathfrak{R}$ die Darstellung von \mathfrak{S} eindeutig bestimmt ist; dabei wird sich herausstellen, daß die $D''(\alpha)$ nur eine Wiederholung der $D(\beta)$ sind, daß es also zu jedem $D''(\alpha)$ ein β so gibt, daß $D(\beta) = D''(\alpha)$. Zuerst zeigen wir:

Satz 6. Die Darstellung von \mathfrak{S} durch die Gruppe der Permutationen $P_{\alpha}(\mu)$ der Räume A_{μ} ist durch die Angabe der Untergruppe \mathfrak{R} von \mathfrak{S} eindeutig bestimmt.

Beweis. Der Satz beruht auf der leicht zu erweisenden Transitivität der Permutationsgruppe, wir zeigen ihn durch Rechnung. Aus (6) folgt, wenn man von der Translation absieht:

$$(16) \quad \alpha_{\mu} u_{1\varrho} = u_{\mu\varrho} \quad (\mu = 1, 2, \dots, d).$$

Dies bedeutet

$$\alpha_{\mu} A_1 = A_{\mu},$$

also

$$P_{\alpha_{\mu}}(1) = \mu.$$

Da nun für $\beta \in \mathfrak{R}$

$$\beta A_1 = A_1,$$

so ist $\alpha_\mu \mathfrak{R}$ der Komplex der Elemente von \mathfrak{H} , die A_1 in A_μ transformieren. Die $\alpha_\mu \mathfrak{R}$ bilden also die d Linksnebenklassen zu \mathfrak{R} in \mathfrak{H} . Nun werden aber die Räume A_μ durch ein Element α von \mathfrak{H} genau so permutiert, wie die $\alpha_\mu \mathfrak{R}$ bei Linksmultiplikation mit α . Sei nämlich

$$\alpha \alpha_\mu \mathfrak{R} = \alpha_\nu \mathfrak{R},$$

also

$$\alpha \alpha_\mu = \alpha_\nu \beta \quad \text{mit} \quad \beta \subset \mathfrak{R},$$

so ist

$$\alpha A_\mu = \alpha \alpha_\mu A_1 = \alpha_\nu \beta A_1 = \alpha_\nu A_1 = A_\nu.$$

Da nun die Zerspaltung von \mathfrak{H} in Linksnebenklassen von \mathfrak{R} (bis auf die Reihenfolge) eindeutig ist, so folgt die behauptete Eindeutigkeit der Permutationsdarstellung.

Berechnung der $D^\mu(\alpha)$. Der Vergleich von (16) mit (10) liefert zunächst

$$(17) \quad D^1(\alpha_\mu) = I,$$

wobei I die s -reihige Einheitsmatrix ist. Ersetzen wir jetzt in (15) β durch α , α durch α_μ und μ durch 1, wobei $P_\alpha(\mu) = P_{\alpha_\mu}(1) = \mu$ wird, so erhalten wir bei Beachtung von (17)

$$(18) \quad D(\alpha \alpha_\mu) = D'(\alpha \alpha_\mu) = D^\mu(\alpha).$$

Damit haben wir die $D^\mu(\alpha)$ durch die $D(\alpha) = D^1(\alpha)$ ausgedrückt. Wenden wir dies auf (15) an, so erhalten wir, wenn wir noch α und β vertauschen,

$$(19) \quad D(\alpha \beta \alpha_\mu) = D(\alpha \alpha_{P_\beta(\mu)}) D(\beta \alpha_\mu).$$

Setzen wir zur Abkürzung

$$(20) \quad \alpha_{P_\beta(1)} = \alpha_\beta$$

und beachten wir, daß $\alpha_1 = \varepsilon$, so ergibt sich für $\mu = 1$

$$(21) \quad D(\alpha \beta) = D(\alpha \alpha_\beta) D(\beta).$$

Da $P_\beta(1) = 1$ für $\beta \subset \mathfrak{R}$, so gilt insbesondere

$$(22) \quad D(\alpha \beta) = D(\alpha) D(\beta) \quad [\beta \subset \mathfrak{R}, \alpha \text{ beliebig!}].$$

Da sich nun jedes Element α von \mathfrak{H} eindeutig in der Form $\alpha_\mu \beta$ mit $\beta \subset \mathfrak{R}$ darstellen läßt, so bestimmen sich sämtliche $D(\alpha)$ aus den $D(\beta)$ mit $\beta \subset \mathfrak{R}$ vermöge

$$(23) \quad D(\alpha) = D(\alpha_\mu \beta) = D(\beta) \quad [\alpha = \alpha_\mu \beta].$$

Bei passender Basiswahl im Raume \mathfrak{S} wird die Darstellung von \mathfrak{H} auf die nur von \mathfrak{R} und $D(\beta)$ abhängige Form (10) gebracht, die demnach durch die innere Struktur von \mathfrak{H} allein bedingt ist, nicht durch die Lagerung von \mathfrak{H} innerhalb \mathfrak{G} .

Aber auch die Darstellung (8) von \mathfrak{G} ist damit bestimmt, wenn man noch die Funktionen $\lambda_r(t)$ kennt. Diese leiten sich nach (5) aus einer einzigen von ihnen ab, nämlich aus

$$(24) \quad \lambda(t) = \lambda_1(t) e^{\frac{2\pi i}{1} \sum_{r=1}^3 \varphi_r n_r} = e^{2\pi i (\varphi, t)}$$

$\lambda(t)$ ist charakterisiert durch einen Vektor φ , den wir uns modulo 1 reduziert denken können, so daß die φ_r positive reelle Zahlen < 1 sind. Der Vektor φ wird durch die Elemente α von \mathfrak{H} kontragredient zu den Translationen transformiert. Wir behaupten nun

Satz 7. *Der Vektor φ ist unter den Elementen β von \mathfrak{R} kontragredient invariant modulo 1.*

Beweis. Es ist

$$(25) \quad \lambda(t) = \lambda(\beta^{-1} t) \quad \text{für } \beta \in \mathfrak{R},$$

also gilt

$$(\varphi, t) \equiv (\varphi, \beta^{-1} t) \quad \text{modulo 1,}$$

oder, wenn wir die Transponierte von β^{-1} mit $(\beta^{-1})'$ bezeichnen:

$$(\varphi, t) \equiv (\varphi^{(\beta^{-1})'}, t) \quad \text{modulo 1.}$$

Dies gilt für alle t , woraus folgt

$$(26) \quad \varphi \equiv \varphi^{(\beta^{-1})'} \quad \text{modulo 1.}$$

Als Ergebnis unserer bisherigen Untersuchung erhalten wir:

Satz 8. *Zu jeder Darstellung von \mathfrak{G} in \mathfrak{S} (und in solche zerfällt nach Satz 3 jede Darstellung von \mathfrak{G}) gehört ein reeller, modulo 1 bestimmter Vektor φ und eine Darstellung $D(\beta)$ derjenigen Untergruppe \mathfrak{R} von \mathfrak{H} , die den Vektor φ modulo 1 kontragredient invariant läßt. Bei geeigneter Basiswahl in \mathfrak{S} läßt sich die Darstellung von \mathfrak{G} auf die Form \mathfrak{S}*

$$(8) \quad t \alpha u_{\mu \nu} = \lambda_r(t) \sum_{\tau} u_{\tau \tau} D_{\tau \tau}^{\mu}(\alpha), \quad (P_{\alpha}(\nu) = \nu)$$

bringen, wobei die $D_{\tau \tau}^{\mu}(\alpha)$ und $P_{\alpha}(\nu)$ durch \mathfrak{R} , $D(\beta)$ und die Struktur von \mathfrak{H} , die $\lambda_r(t)$ durch φ und die arithmetische Kristallklasse von \mathfrak{G} bestimmt sind, das heißt dadurch, wie \mathfrak{H} auf die Translationsgruppe Γ wirkt.

§ 2.

Wir zeigen jetzt, daß jeder Ausdruck, der in der Form (8) angesetzt wird, auch umgekehrt eine Darstellung von \mathfrak{G} liefert. Genauer bedeutet dies folgendes: Wir gehen aus von einem beliebigen Vektor mit reellen Komponenten, der durch die Elemente von \mathfrak{H} kontragredient zu den Translationen von Γ transformiert wird. \mathfrak{R} sei die Untergruppe derjenigen Elemente β

von \mathfrak{H} , die φ modulo 1 kontragredient invariant lassen (\mathfrak{R} kann auch mit \mathfrak{H} identisch sein oder auch nur aus dem Einselement ε bestehen). $D(\beta)$ sei endlich eine beliebige s -dimensionale Darstellung von \mathfrak{R} . \mathfrak{R} habe den Index d . Wir zerlegen \mathfrak{H} nach \mathfrak{R} in die Linksnebenklassen α_μ ($\alpha_1 = \varepsilon$), die bei Linksmultiplikation mit α eine Permutation $P_\alpha(\mu)$ erfahren. Aus φ bilden wir nun die Funktion

$$(24) \quad \lambda(t) = e^{2\pi i \langle \eta, t \rangle}$$

und daraus die Funktionen

$$\lambda_\mu(t) = \lambda(t^{\alpha_\mu^{-1}}).$$

Für sie gilt die Formel (7); denn sei $P_\alpha(\mu) = \nu$, also $\alpha\alpha_\mu\mathfrak{R} = \alpha_\nu\mathfrak{R}$, so hat man

$$\lambda_{P_\alpha(\mu)}(t) = \lambda(t^{\alpha_\nu^{-1}}).$$

Andererseits ist

$$(\varphi, t^{\alpha_\mu^{-1}}) = (\varphi, t^{\alpha_\nu\alpha_\mu^{-1}}) = (\varphi, [t^{\alpha_\mu^{-1}}]^{\alpha_\nu}) = (\varphi^{(t^{-1})^\nu}, t^{\alpha_\nu^{-1}});$$

wegen der kontragredienten Invarianz von φ ist aber

$$(\varphi^{(t^{-1})^\nu}, t^{\alpha_\nu^{-1}}) \equiv (\varphi, t^{\alpha_\nu^{-1}}) \pmod{1},$$

und deshalb folgt

$$(7) \quad \lambda_{P_\alpha(\mu)}(t) = \lambda_\mu(t^{\alpha_\nu^{-1}}).$$

Wir bilden sodann mit Hilfe der $D(\beta)$ ($\beta \in \mathfrak{R}$) die s -reihigen Matrizes $(D_{\tau\varrho}^\mu(\alpha))$ für alle $\alpha \in \mathfrak{H}$, indem wir setzen:

$$(23) \quad D(\alpha_\mu\beta) = D(\beta)$$

und

$$(18) \quad D^\mu(\alpha) = D(\alpha\alpha_\mu).$$

Schließlich definieren wir zu $t\alpha$ aus \mathfrak{G} eine lineare Transformation eines Vektorraumes $\mathfrak{S} = (u_{\mu\varrho})$ ($\mu = 1, 2, \dots, d$, $\varrho = 1, 2, \dots, s$) durch

$$(8) \quad t\alpha u_{\mu\varrho} = \lambda_\mu(t) \sum_\tau u_{\tau\tau} D_{\tau\varrho}^\mu(\alpha) \quad (P_\alpha(\mu) = \nu)$$

und zeigen, daß dies eine Darstellung von \mathfrak{G} ist, daß also für zwei Elemente $s\beta$ und $t\alpha$ aus \mathfrak{G} gilt:

$$(s\beta t\alpha) u_{\mu\varrho} = s\beta(t\alpha u_{\mu\varrho}).$$

Sei $P_\beta(\nu) = \kappa$, dann haben wir

$$s\beta(t\alpha u_{\mu\varrho}) = \lambda_\kappa(s) \lambda_\mu(t) \sum_\omega \sum_\tau u_{\omega\tau} D_{\omega\tau}^\nu(\beta) D_{\tau\varrho}^\mu(\alpha).$$

Andererseits ist

$$s\beta t\alpha = s\beta^t\beta\alpha$$

und

$$P_{\beta\alpha}(\mu) = P_\beta(P_\alpha(\mu)) = P_\beta(\nu) = \kappa,$$

also

$$(s\beta t\alpha)u_{\mu\epsilon} = \lambda_x(st^3) \sum_{\alpha} u_{\alpha} D_{\alpha\epsilon}^u(\beta\alpha).$$

Da aber, wie auf S. 201 durch Anwendung von (7) gezeigt wurde,

$$\lambda_x(st^3) = \lambda_x(s)\lambda_x(t)$$

ist, so bleibt uns nur noch zu zeigen, daß

$$(15) \quad D^u(\beta\alpha) = D^{P_{\alpha}(u)}(\beta) D^u(\alpha)$$

ist. Dies kann man aber nach (18) schreiben:

$$(19) \quad D(\alpha\beta\alpha_{\mu}) = D(\alpha\alpha_{P_{\beta}(u)})D(\beta\alpha_{\mu}).$$

Wir haben also zu zeigen, daß dies eine Folge von

$$(23) \quad D(\alpha_{\mu}\beta) = D(\beta)$$

ist.

Sei α ein beliebiges Element aus \mathfrak{H} , β und γ Elemente aus \mathfrak{R} , dann folgt aus (23)

$$D(\alpha\beta) = D(\alpha, \gamma\beta) = D(\alpha, (\gamma\beta)) = D(\gamma\beta) = D(\gamma) D(\beta)$$

aber

$$D(\alpha) = D(\alpha, \gamma) = D(\gamma).$$

Es ist also

$$(22) \quad D(\alpha\beta) = D(\alpha) D(\beta) \quad [\beta \in \mathfrak{R}, \alpha \text{ beliebig}]$$

eine Folgerung von (23).

Sei jetzt $\alpha = \alpha, \beta$ und $\delta = \alpha_{\mu}\gamma$, so ist

$$D(\alpha\delta) = D(\alpha, \beta\alpha_{\mu}\gamma) = D(\alpha, \beta\alpha_{\mu}) D(\gamma) = D(\alpha, \beta\alpha_{\mu}) D(\alpha_{\mu}\gamma).$$

Es ist aber, wenn wir wieder zur Abkürzung

$$\alpha_{P_{\beta}(1)} = \alpha_{\delta}$$

setzen,

$$\alpha_{\delta} = \alpha_{P_{\alpha_{\mu}\gamma}(1)} = \alpha_{P_{\alpha_{\mu}}(1)} = \alpha_{\mu}.$$

Also folgt aus (22) die Formel (21) für $\beta = \delta$

$$D(\alpha\delta) = D(\alpha, \beta\alpha_{\delta}) D(\alpha_{\mu}\gamma) = D(\alpha\alpha_{\delta}) D(\delta).$$

Hieraus folgt endlich, wenn wir $\delta = \beta\alpha_{\mu}$ setzen:

$$D(\alpha\beta\alpha_{\mu}) = D(\alpha\alpha_{P_{\beta\alpha_{\mu}}(1)}) D(\beta\alpha_{\mu}) = D(\alpha\alpha_{P_{\beta}(u)}) D(\beta\alpha_{\mu}).$$

Dies ist aber Gleichung (19), aus der folgt, daß (8) wirklich eine Darstellung von \mathfrak{G} liefert.

§ 3.

Wir suchen unter den so erhaltenen Darstellungen von \mathfrak{G} die irreduziblen und die inäquivalenten. Es gilt zunächst

Satz 9. *Dann und nur dann, wenn die Darstellung $D(\beta)$ von \mathfrak{R} irreduzibel ist, ist auch die zugehörige Darstellung (8) von \mathfrak{G} irreduzibel.*

Beweis. Zerfällt die Darstellung $D(\beta)$, so zerfällt auch die Darstellung von \mathfrak{G} , wie man sofort einsieht, wenn man die Matrix (9) betrachtet und daran denkt, daß in den einzelnen Kästchen lauter Matrizen $D(\beta)$ stehen.

Sei nun umgekehrt $D(\beta)$ irreduzibel, dann zeigen wir, daß die Darstellung von \mathfrak{G} in \mathfrak{S} irreduzibel ist, daß also eine Matrix A , die mit den Darstellungsmatrizen $D(\alpha)$ vertauschbar ist, von der Form λI ist, wobei λ eine komplexe Zahl und I die Einheitsmatrix ist. Zunächst muß A mit den Matrizen, die I zugeordnet sind, vertauschbar sein, diese sind aber von der Form

$$\vartheta(t) = \begin{pmatrix} \lambda_1 I & & & \\ & \lambda_2 I & & \\ & & \ddots & \\ & & & \lambda_d I \end{pmatrix},$$

wobei die λ_v alle verschieden sind. A hat also die Form

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_d \end{pmatrix}.$$

Wir behaupten ferner, daß alle A_v einander gleich sind. Da nämlich $D(\alpha_v) = I$, so haben wir

$$\vartheta(\alpha, \alpha_v) = \begin{pmatrix} 0 \\ \vdots \\ \lambda_v(\alpha_v) I & 0 \cdots 0 \\ \vdots \\ 0 \end{pmatrix},$$

$$\vartheta(\alpha, \alpha_v) A = \begin{pmatrix} 0 \\ \vdots \\ \lambda_v(\alpha_v) A_v & 0 \cdots 0 \\ \vdots \\ 0 \end{pmatrix}, \quad A \vartheta(\alpha, \alpha_v) = \begin{pmatrix} 0 \\ \vdots \\ \lambda_v(\alpha_v) A_v & 0 \cdots 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Aus der Vertauschbarkeit von A mit $\vartheta(a, \alpha_r)$ folgt also $A_1 = A_r$. Schließlich zeigen wir noch, daß A_1 von der Form λI ist; sei nämlich $\beta \in \mathfrak{R}$, so ist

$$\vartheta(\beta) = \begin{pmatrix} \lambda_1(\beta) D(\beta) \cdots 0 \\ \vdots \\ 0 \end{pmatrix},$$

$$\vartheta(\beta) A = \begin{pmatrix} \lambda_1(\beta) D(\beta) A_1 \cdots 0 \\ \vdots \\ 0 \end{pmatrix}, \quad A \vartheta(\beta) = \begin{pmatrix} \lambda_1(\beta) A_1 D(\beta) \cdots 0 \\ \vdots \\ 0 \end{pmatrix}.$$

A_1 muß also mit dem irreduziblen System der $D(\beta)$ vertauschbar sein, hat also die Form $A_1 = \lambda I$, q. e. d.

Wir wissen jetzt, wie wir zu den irreduziblen Darstellungen von \mathfrak{G} gelangen können. Wie finden wir nur die inäquivalenten unter ihnen? Betrachten wir zunächst eine Darstellung $\vartheta(t\alpha)$ in einem Raum \mathfrak{S} , der, wie wir wissen, unter Γ in die d Räume A_r zerfällt. Jeder dieser Räume gehört zu einem Eigenwert $\lambda_r(t) = e^{2\pi i \langle \varphi_r, t \rangle}$, also gehört zu jedem A_r ein kontragredienter Vektor

$$\varphi_r = \varphi^{(\alpha_r^{-1})}.$$

$\varphi = \varphi_1$, und damit auch A_1 , sind unter $\mathfrak{R} = \mathfrak{R}_1$ invariant (wenn wir von den Translationen absehen). \mathfrak{R} erfährt in A_1 die Darstellung $\beta \rightarrow D(\beta)$. A_r und φ_r sind aber invariant unter der zu \mathfrak{R} konjugierten Untergruppe $\mathfrak{R}_r = \alpha_r \mathfrak{R} \alpha_r^{-1}$ von \mathfrak{H} , die zu \mathfrak{R} isomorph ist und deshalb dieselben irreduziblen Darstellungen wie \mathfrak{R} besitzt. \mathfrak{R}_r erfährt in A_r die Darstellung

$$\alpha_r \beta \alpha_r^{-1} \rightarrow D^r(\alpha_r \beta \alpha_r^{-1}) = D(\alpha_r \beta \alpha_r^{-1} \alpha_r) = D(\alpha_r \beta) = D(\beta).$$

Die d Vektoren φ_r sind alle verschieden. Man erhält sie aus einem beliebigen unter ihnen, wenn man ihn mit Hilfe aller Elemente von \mathfrak{H} kontragredient transformiert. Wir können nun ausgehend von φ , eine Darstellung $\vartheta^r(t\alpha)$ von \mathfrak{G} in der Form (8) bilden, indem wir eine Darstellung von \mathfrak{R}_r wählen, die der Darstellung $D(\beta)$ von \mathfrak{R} äquivalent ist. Diese Darstellung $\vartheta^r(t\alpha)$ wird, so behaupten wir, der Darstellung $\vartheta(t\alpha)$ äquivalent sein. Denn wir werden wieder genau d Räume A_μ^r erhalten, von denen einer zu $\varphi_1 = \varphi$ gehören muß. Ihn wird gerade die Untergruppe \mathfrak{R} , die zu \mathfrak{R}_r isomorph ist, invariant lassen und in ihm eine zu $D(\beta)$ äquivalente Darstellung erfahren. Durch gleichzeitigen Basiswechsel in allen Räumen A_μ^r erreichen wir, daß \mathfrak{R} gerade die Darstellung $D(\beta)$ erfährt, und jetzt können wir, wie im 1. Teil gezeigt wurde, die Darstellung auf die eindeutig durch φ , \mathfrak{R} und $D(\beta)$ bestimmte Form $D(t\alpha)$ bringen.

Wir wollen jetzt die Vektoren φ_r konjugiert in bezug auf \mathfrak{H} nennen und uns einen Bereich F des Einheitswürfels vorstellen, derart, daß zu jedem

kontragredienten Vektor ψ ein in bezug auf \mathfrak{H} konjugierter φ in F liegt, daß es also zu ψ ein $\alpha \in \mathfrak{H}$ gibt derart, daß

$$\psi\alpha^{-1} \equiv \varphi \pmod{1} \quad \text{mit } \varphi \in F$$

und daß ferner in F keine zwei Vektoren konjugiert sind. Zu jedem $\varphi \in F$ gibt es dann eine Untergruppe $\mathfrak{R}_\varphi \subset \mathfrak{H}$, die φ invariant läßt und endlich viele irreduzible Darstellungen von \mathfrak{R}_φ . Wir behaupten dann:

Satz 10. *Zu jedem $\varphi \in F$ gibt es genau so viele irreduzible Darstellungen von \mathfrak{G} , als es irreduzible und inäquivalente Darstellungen von \mathfrak{R}_φ gibt. Alle diese Darstellungen zusammen bilden ein vollständiges System Σ von irreduziblen und inäquivalenten Darstellungen von \mathfrak{G} .*

Beweis. Wir haben bereits gezeigt, daß es zu jeder irreduziblen Darstellung von \mathfrak{G} eine äquivalente in Σ gibt. Gesetzt nun, in Σ seien zwei verschiedene äquivalente Darstellungen ϑ und $\bar{\vartheta}$ von \mathfrak{G} , die zu den Vektoren ψ und $\bar{\psi}$ gehören sollen. Dann zerfallen die Darstellungsräume γ und $\bar{\gamma}$ unter I' in invariante Teilräume Λ_i und $\bar{\Lambda}_i$, die in irgendeiner Reihenfolge in bezug auf I' einander operatorisomorph sein müssen. Zu Λ_1 sei etwa $\bar{\Lambda}_\mu$ operatorisomorph, was bedeutet, daß Λ_1 und $\bar{\Lambda}_\mu$ zum selben Eigenwert gehören, also

$$\varphi = \varphi_1 \equiv \bar{\varphi}_\mu \pmod{1}.$$

Nun ist $\bar{\varphi}$ zu $\bar{\varphi}_\mu$ konjugiert. Also sind auch φ und $\bar{\varphi}$ einander konjugiert, was nach der Konstruktion von F nur möglich ist, wenn $\varphi = \bar{\varphi}$. Λ_1 und $\bar{\Lambda}_1$ sind also invariant unter derselben Untergruppe \mathfrak{R} . Da nun Λ_1 und $\bar{\Lambda}_1$ äquivalent sind, so erfährt auch \mathfrak{R} in Λ_1 und $\bar{\Lambda}_1$ äquivalente Darstellungen. Σ enthält aber zu jedem φ und \mathfrak{R} nur so viele Darstellungen, als es irreduzible und inäquivalente Darstellungen von \mathfrak{R} gibt. ϑ und $\bar{\vartheta}$ können also im Widerspruch zur Voraussetzung nicht zwei verschiedene Darstellungen sein.

§ 4.

In der Darstellungstheorie der endlichen Gruppen erhalten wir eine Übersicht über alle irreduziblen Darstellungen durch den Satz, daß jede irreduzible Darstellung vom Grade n genau n -mal in der regulären Darstellung enthalten ist. Die reguläre Darstellung einer endlichen Gruppe \mathfrak{H} kann man erhalten, indem man den h Elementen α , von \mathfrak{H} die h Basisvektoren α_μ eines Vektorraumes \mathfrak{S} zuordnet. Bei Linksmultiplikation mit einem Element α_μ von \mathfrak{H} erfahren die α , eine Permutation, und man ordnet dem Element α_μ die lineare Transformation von \mathfrak{S} zu, die der entsprechenden Permutation der α , entspricht.

Es gelingt nun, diesen Satz auf unsere Theorie der beschränkten Darstellungen der Raumgruppen zu übertragen, indem wir zu jedem Vektor φ

die „reguläre Darstellung von \mathfrak{G} zum Vektor φ “ definieren. Sie entsteht aus einer Erweiterung der regulären Darstellung von \mathfrak{H} dadurch, daß wir die Wirkung von Γ auf die Basisvektoren u_r festsetzen, und zwar gemäß den Gleichungen auf S. 197.

$$t u_r = \lambda_r(t) u_r = e^{2\pi i(\varphi, t^{-1})} u_r.$$

Die Wirkung eines beliebigen Elements von \mathfrak{G} auf die h Basisvektoren ist damit klar, und aus den früheren Rechnungen geht hervor, daß wir es wirklich mit einer Darstellung von \mathfrak{G} zu tun haben. Unter der *Zähligkeit* von φ wollen wir den Index der Untergruppe \mathfrak{R}_φ von \mathfrak{H} verstehen. Ein Vektor, der nur unter s invariant ist, ist also h -zählig. Besitzt ein Vektor φ in bezug auf \mathfrak{H} d konjugierte, so ist er d -zählig. d ist ein Teiler von h . Nachdem wir diese Begriffe eingeführt haben, können wir aussprechen:

Satz 11. Jede irreduzible Darstellung vom Grade n von \mathfrak{G} zum Vektor φ der Zähligkeit d ist schon in der regulären Darstellung von \mathfrak{G} zum Vektor φ enthalten, und zwar genau $\frac{n}{d}$ -mal.

Beweis. Sei φ invariant unter \mathfrak{R} , dann enthält der Raum der regulären Darstellung einen unter Γ und \mathfrak{R} invarianten Teilraum A_1 der Dimension $\frac{h}{d}$, in dem die Untergruppe \mathfrak{R} selbst die reguläre Darstellung erfährt. Diese zerfällt, wie wir wissen, in irreduzible Bestandteile derart, daß jede Darstellung von \mathfrak{R} so oft vorkommt, als ihr Grad beträgt. Ein Blick auf die Matrix (9) von S. 200 lehrt aber, daß dies eine Zerfällung von \mathfrak{G} in irreduzible Bestandteile bedeutet. Satz 10 lehrt uns, daß wir auf diese Weise alle irreduziblen Darstellungen von \mathfrak{G} zum Vektor φ erhalten. Ist eine von ihnen n -dimensional, so entspricht ihr eine n/d -dimensionale irreduzible Darstellung von \mathfrak{R} in A_1 , womit alles gezeigt ist.

§ 5.

Die Punktgruppe \mathfrak{H} von \mathfrak{G} läßt das von Γ erzeugte Gitter invariant, das durch die Matrix γ charakterisiert werden kann, die den Übergang von Gitterkoordinaten zu kartesischen Koordinaten vollzieht. Die arithmetische Kristallklasse C von \mathfrak{G} wird definiert durch die Darstellung von \mathfrak{H} in Gitterkoordinaten. Wir wollen nun die Elemente von \mathfrak{H} , die uns durch ihre Matrizes α in Gitterkoordinaten gegeben seien, auf das reziproke Gitter beziehen, das bekanntlich durch die Matrix $\gamma^* = (\gamma^{-1})'$ charakterisiert ist. Auf kartesische Koordinaten bezogen bilden die Elemente von \mathfrak{H} orthogonale Matrizes α_k , für die

$$\alpha_k^* = (\alpha_k^{-1})' = \bar{\alpha}_k$$

gilt. Im reziproken Gitter werden sie also durch die Matrizes

$$\gamma^{*-1} \alpha_k \gamma^* = \gamma^{*-1} \alpha_k^* \gamma^* = (\gamma^{-1} \alpha_k \gamma)^* = \alpha^*$$

dargestellt. Die α^* sind ebenfalls unimodular und ganzzahlig und definieren die zu C reziproke arithmetische Kristallklasse C^* . Da der Übergang von C zu C^* durch die lineare Transformation $\gamma^{*-1} \gamma = \gamma' \gamma$ geschieht, so gehören C und C^* immer zur selben geometrischen Kristallklasse. In 47 Fällen läßt sich der Übergang aber auch ganzzahlig unimodular erreichen, dann sind die Klassen zu sich selbst reziprok. Die übrigen 26 arithmetischen Kristallklassen bilden 13 zueinander reziproke Paare, wie man durch Rechnung feststellen kann.

Sei jetzt \mathfrak{G}^* die symmorphie Raumgruppe von C^* , also die Raumgruppe, die von den Elementen von \mathfrak{H} selbst und von der Translationsgruppe I^* des reziproken Gitters erzeugt wird, dann sehen wir, daß der in § 3 definierte Bereich F eine anschauliche Deutung als Fundamentalbereich von \mathfrak{G}^* zuläßt. Wir deuten nämlich die Koordinaten eines kontragredienten Vektors φ als die auf das reziproke Gitter bezogenen Koordinaten eines Punktes φ . Dann entsprechen konjugierten Vektoren im Sinne des § 3 gerade unter \mathfrak{G}^* konjugierte Punkte. Wir können F als konvexes Polyeder im primitiven Parallelepipied des reziproken Gitters annehmen. Mit $\alpha^* F$ bezeichnen wir das Polyeder, in das F durch α^* übergeht, sowie alle Polyeder, die aus $\alpha^* F$ durch Translationen von I^* entstehen. Dann füllen die $\alpha^* F$ den ganzen Raum lückenlos aus. F selbst ist von allen Seiten mit Polyedern $\alpha^* F$ umgeben, die an seine Seitenflächen, Kanten und Ecken anstoßen. Wir kommen so zu einer anschaulichen geometrischen Übersicht über alle irreduziblen Darstellungen von \mathfrak{G} .

Ein innerer Punkt von F wird außer von ε von keinem weiteren Element aus \mathfrak{H} modulo I^* invariant gelassen. Er ist also h -zählig und zu ihm gehört eine einzige irreduzible Darstellung: die reguläre Darstellung. Entsprechend den 3 Freiheitsgraden des inneren Punktes von F erhalten wir eine Schar von ∞^3 Darstellungen vom Grade h .

Ein Punkt einer Seitenfläche σ von F , der nicht auf einer Kante liegt, gehört außer F noch genau einem $\alpha^* F$ an, jedoch kann er nur dann unter α^* invariant sein, wenn $\alpha^* F$ mit der Fläche $\alpha^* \sigma$ an σ anstößt. Ist dies der Fall, so bedeutet α^* auf σ angewandt (immer unter Vernachlässigung von Translationen aus I^*) entweder eine Drehung um den Mittelpunkt, eine Spiegelung an einer Symmetrieachse oder die Identität. Es bleibt also entweder kein Punkt, ein einziger Punkt, eine Strecke oder das ganze Innere von σ invariant, und zwar genau unter α^* und ε , also unter einer Untergruppe \mathfrak{R} der Ordnung 2 von \mathfrak{H} . Entsprechend den zwei irreduziblen Darstellungen dieser Untergruppe erhalten wir also entweder gar keine Darstellung, genau 2 Darstellungen,

2 Scharen von ∞^1 Darstellungen oder 2 Scharen von ∞^2 Darstellungen vom Grade $\frac{h}{2}$ für die invarianten inneren Punkte von σ . Zu den übrigen Punkten im Innern von σ gehört die reguläre Darstellung vom Grade h .

Analoges gilt für die Punkte auf den Kanten von F , nur ist hier die Untergruppe \mathfrak{R} im allgemeinen komplizierter. Eine ausgezeichnete Rolle können die Kantenmittelpunkte spielen. Zu ihnen und zu den Ecken von F gehören einzelne Darstellungen, deren Form, Grad und Zahl sich aus den Untergruppen \mathfrak{R} von \mathfrak{H} ergibt, unter denen die Kantenmittelpunkte und Ecken festbleiben.

Ein Punkt auf einer Kante K von F hängt von einem Parameter ab, deshalb gehören zu ihm Scharen von ∞^1 Darstellungen. Läßt man nun den Parameter frei variieren, so wandert der Punkt auf eine Kante α^*G des an F anstoßenden α^*F , wobei durchaus nicht $K = G$ sein muß. Wir wollen solche Kanten *zusammengehörig* nennen und analog von zusammengehörigen Flächen sprechen. Die Darstellungen, die zu zusammengehörigen Kanten oder Flächen gehören, lassen sich in eine Schar zusammenfassen, wenn man sich für diesen Fall nicht auf F streng beschränkt. Dies ist wegen des engeren Anschlusses an die bestehenden Berechnungen der Raumgruppen vorteilhaft.

Bevor wir dies im folgenden Paragraphen noch durch ein Beispiel illustrieren, sei abschließend bemerkt, daß die hier für dreidimensionale Raumgruppen entwickelte Theorie auch für beliebige n -dimensionale Raumgruppen gilt. Die Sätze von § 1 bis 4 bleiben darüber hinaus gültig, wenn \mathfrak{H} eine beliebige endliche Gruppe, Γ eine Abelsche Gruppe der Struktur $\mathfrak{C}_{\infty} \dots$ und \mathfrak{G} eine Erweiterung von Γ mit \mathfrak{H} ist.

§ 6.

Zur Berechnung der Darstellungen kann man sich der „Internationalen Tabellen zur Bestimmung von Kristallstrukturen“ [5] bedienen. Ein Hilfsmittel bietet die algebraische Aufstellung der Raumgruppen in kartesischen Koordinaten von F. Seitz [4]. Die arithmetische Kristallklasse zu jeder vorgegebenen Raumgruppe und die zugehörige symmorphe Raumgruppe findet man bei Niggli und Nowacki [6]. Die dieser Arbeit beigelegte Tabelle 1 gestattet zu jeder arithmetischen Klasse die reziproke Klasse, die geometrische Klasse und ihre gruppentheoretische Struktur abzulesen, sowie die Klassengleichung, die Zahl und Grad ihrer irreduziblen Darstellungen angibt. Zur Erläuterung von Tabelle 1 betrachten wir etwa Zeile 10. Das Symbol $C_{3v}(\alpha\delta\epsilon)$ bedeutet: Die arithmetischen Klassen $C_{3v\alpha}$, $C_{3v\beta}$, $C_{3v\epsilon}$ (Bezeichnung nach Niggli und Nowacki [6]) gehören zur geometrischen Klasse C_{3v} . $C_{3v\delta}$ und $C_{3v\epsilon}$ sind zueinander reziprok, während $C_{3v\alpha}$ zu sich selbst reziprok ist.

Die geometrischen Klassen C_{3v} und D_3 haben die Struktur \mathbb{C}_3^D (\mathbb{C} , bedeutet zyklische Gruppe der Ordnung v , \mathbb{C}^D dazugehörige Diedergruppe, T Tetraedergruppe, O Oktaedergruppe); gemäß der Klassengleichung besitzen die Klassen von Zeile 10 zwei eindimensionale und eine zweidimensionale Darstellung.

Tabelle 1. Gruppentheoretische Struktur der Kristallklassen.

Nr.	Arithmetische und geometrische Kristallklassen	Struktur	Klassengleichung
1	C_1	e	$1 \cdot 1^2 = 1$
2	C_2 $C_2(\alpha\beta)$ $C_2(\alpha\delta)$	\mathbb{C}_2	$2 \cdot 1^2 = 2$
3	$C_3(\alpha\delta)$	\mathbb{C}_3	$3 \cdot 1^2 = 3$
4	$C_4(\alpha\beta)$ $S_4(\alpha\beta)$	\mathbb{C}_4	$4 \cdot 1^2 = 4$
5	$C_{3v}(\alpha\delta)$ C_{3h} C_6	\mathbb{C}_6	$6 \cdot 1^2 = 6$
6	$C_{2h}(\alpha\beta)$ $C_{2v}(\alpha\beta\gamma\delta\epsilon)$ $D_2(\alpha\beta\gamma\delta)$	$\mathbb{C}_2 \times \mathbb{C}_2$	$4 \cdot 1^2 = 4$
7	$D_{2h}(\alpha\beta\gamma\delta)$	$\mathbb{C}_2 \times \mathbb{C}_2 \times \mathbb{C}_2$	$8 \cdot 1^2 = 8$
8	$C_{4h}(\alpha\beta)$	$\mathbb{C}_4 \times \mathbb{C}_2$	$8 \cdot 1^2 = 8$
9	C_{6h}	$\mathbb{C}_6 \times \mathbb{C}_2$	$12 \cdot 1^2 = 12$
10	$C_{3v}(\alpha\delta\epsilon)$ $D_3(\alpha\delta\epsilon)$	\mathbb{C}_3^D	$2 \cdot 1^2 + 1 \cdot 2^2 = 6$
11	$C_{4v}(\alpha\beta)$ $D_{2d}(\alpha\beta\gamma\delta)$ $D_4(\alpha\beta)$	\mathbb{C}_4^D	$4 \cdot 1^2 + 1 \cdot 2^2 = 8$
12	$D_{3d}(\alpha\delta\epsilon)$ $D_{3h}(\delta\epsilon)$ C_{6v} D_6	\mathbb{C}_6^D	$4 \cdot 1^2 + 2 \cdot 2^2 = 12$
13	$D_{4h}(\alpha\beta)$	$\mathbb{C}_4^D \times \mathbb{C}_2$	$8 \cdot 1^2 + 2 \cdot 2^2 = 16$
14	D_{6h}	$\mathbb{C}_6^D \times \mathbb{C}_2$	$8 \cdot 1^2 + 4 \cdot 2^2 = 24$
15	$T(\alpha\beta\gamma)$	T	$3 \cdot 1^2 + 1 \cdot 3^2 = 12$
16	$T_d(\alpha\beta\gamma)$ $O(\alpha\beta\gamma\delta\epsilon)$	O	$2 \cdot 1^2 + 1 \cdot 2^2 + 2 \cdot 3^2 = 24$
17	$T_h(\alpha\beta\gamma)$	$T \times \mathbb{C}_2$	$6 \cdot 1^2 + 2 \cdot 3^2 = 24$
18	$O_h(\alpha\beta\gamma)$	$O \times \mathbb{C}_2$	$4 \cdot 1^2 + 2 \cdot 2^2 + 4 \cdot 3^2 = 48$

Als Beispiel betrachten wir die Raumgruppen der hexagonalen Holoedrie, die alle zur selben zu sich selbst reziproken arithmetischen Klasse D_{6h} gehören. D_{6h} , bestehend aus allen euklidischen Bewegungen und Umlegungen, die ein hexagonales Gitter und den Gitterpunkt 0 festlassen, wird erzeugt von einer Drehung ξ der Ordnung 6 um die z -Achse, einer Spiegelung δ an der x - z -Ebene und einer Spiegelung η an der x - y -Ebene eines kartesischen Koordinatensystems. Die Gruppe \mathbb{G}^* ist hier D_{6h}^1 . Wir legen das reziproke Gitter so fest wie das Gitter von D_{6h}^1 in den „Internationalen Tabellen“. Dann wird in der dort angewandten kristallographischen Schreibweise

$$\xi^* = [x - yxz], \quad \delta^* = [yzx], \quad \eta^* = [xy\bar{z}],$$

was heißen soll, daß der Punkt (x, y, z) durch ξ^* , δ^* , η^* in die Punkte $(x - y, x, z)$, (y, x, z) und $(x, y, -z)$ übergeführt wird. Der Fundamental-

bereich F ist ein dreikantiges Prisma mit den 6 Ecken

$$\begin{aligned} a_1 &= (0 \ 0 \ 0) & b_1 &= (\tfrac{1}{2} \ 0 \ 0) & c_1 &= (\tfrac{2}{3} \ \tfrac{1}{3} \ 0) \\ a_2 &= (0 \ 0 \ \tfrac{1}{2}) & b_2 &= (\tfrac{1}{2} \ 0 \ \tfrac{1}{2}) & c_2 &= (\tfrac{2}{3} \ \tfrac{1}{3} \ \tfrac{1}{2}). \end{aligned}$$

F hat $2 \cdot 3 + 3 = 9$ Kanten, nämlich

$$\begin{aligned} A &= a_1 a_2 & B &= b_1 b_2 & C &= c_1 c_2 \\ A_1 &= b_1 c_1 & B_1 &= a_1 c_1 & C_1 &= a_1 b_1 \\ A_2 &= b_2 c_2 & B_2 &= a_2 c_2 & C_2 &= a_2 b_2; \end{aligned}$$

dabei sind A_1 und B_1 sowie A_2 und B_2 zusammengehörige Kanten. F hat ferner 5 Seitenflächen, nämlich

$$\begin{aligned} \alpha &= b_1 c_1 b_2 c_2 & \beta &= a_1 c_1 a_2 c_2 & \gamma &= a_1 b_1 a_2 b_2 \\ \alpha_1 &= a_1 b_1 c_1 & \alpha_2 &= a_2 b_2 c_2, \end{aligned}$$

Tabelle 2. Die irreduziblen Darstellungen der Raumgruppen der hexagonalen Holoedrie.

Punktlage		Symmetrie R	Index ($D_{6h}:R$)	Zahl der Darstellungen vom Grade							
				1	2	3	4	6	8	12	24
a_1	0 0 0	$D_{6h} \cong \mathbb{C}_6^D \times \mathbb{C}_2$	1	8	4						
b_1	$\frac{1}{2}$ 0 0	$D_{3h} \cong \mathbb{C}_3 \times \mathbb{C}_2 \times \mathbb{C}_2$	3			8					
c_1	$\frac{2}{3}$ $\frac{1}{3}$ 0	$D_{3h} \cong \mathbb{C}_6^D$	2		4		2				
a_2	0 0 $\frac{1}{2}$	$D_{6h} \cong \mathbb{C}_6^D \times \mathbb{C}_2$	1	8	4						
b_2	$\frac{1}{2}$ 0 $\frac{1}{2}$	$D_{3h} \cong \mathbb{C}_3 \times \mathbb{C}_2 \times \mathbb{C}_2$	3			8					
c_2	$\frac{2}{3}$ $\frac{1}{3}$ $\frac{1}{2}$	$D_{3h} \cong \mathbb{C}_6^D$	2		4		2				
A	0 0 z	$C_{6v} \cong \mathbb{C}_6^D$	2		$4 \cdot \infty^1$		$2 \cdot \infty^1$				
B	$\frac{1}{2}$ 0 z	$C_{2v} \cong \mathbb{C}_2 \times \mathbb{C}_2$	6					$4 \cdot \infty^1$			
C	$\frac{2}{3}$ $\frac{1}{3}$ z	$C_{3v} \cong \mathbb{C}_3^D$	4				$2 \cdot \infty^1$		$1 \cdot \infty^1$		
A_1 B_1	$2x \ x \ 0$	$C_{2v} \cong \mathbb{C}_2 \times \mathbb{C}_2$	6					$4 \cdot \infty^1$			
C_1	$x \ 0 \ 0$	$C_{2v} \cong \mathbb{C}_2 \times \mathbb{C}_2$	6								
A_2 B_2	$2x \ x \ \frac{1}{2}$	$C_{3v} \cong \mathbb{C}_2 \times \mathbb{C}_2$	6					$4 \cdot \infty^1$			
C_2	$x \ 0 \ \frac{1}{2}$	$C_{2v} \cong \mathbb{C}_2 \times \mathbb{C}_2$	6					$4 \cdot \infty^1$			
α β	$2x \ x \ z$	$C_s \cong \mathbb{C}_2$	12							$2 \cdot \infty^2$	
γ	$x \ 0 \ z$	$C_s \cong \mathbb{C}_2$	12							$2 \cdot \infty^2$	
α_1	$x \ y \ 0$	$C_s \cong \mathbb{C}_2$	12							$2 \cdot \infty^2$	
α_2	$x \ y \ \frac{1}{2}$	$C_s \cong \mathbb{C}_2$	12							$2 \cdot \infty^2$	
F	$x \ y \ z$	$C_1 \cong \mathbb{C}$	24								$1 \cdot \infty^3$

α und β sind zusammengehörig. F wird so von den α^*F umgeben, daß immer α^*F mit $\alpha^*\sigma$ an die Fläche σ und mit α^*K an die Kante K stößt, wobei alle Punkte von σ und K invariant bleiben.

Die in Ecken, Flächen und Kanten herrschenden Symmetrien, d. h. die Untergruppen \mathfrak{R} , kann man aus den „Internationalen Tabellen“ ohne weiteres ablesen. Tabelle 2 gibt die vollständige Übersicht über die irreduziblen Darstellungen einer Raumgruppe \mathfrak{G} der Klasse D_{6h} , wie sie sich mit Hilfe von Tabelle 1 ohne weitere Rechnung herstellen läßt.

Um noch zu zeigen, wie sich jetzt die Darstellungen explizit aufstellen lassen, behandeln wir die zur Kante C gehörigen Darstellungen. \mathfrak{R} ist C_{3v} , erzeugt von ξ^2 und $\xi^3\delta$. Aus der Zerlegung von \mathfrak{H} in Nebenklassen nach \mathfrak{R}

$$\mathfrak{H} = \mathfrak{R} + \xi\mathfrak{R} + \eta\mathfrak{R} + \xi\eta\mathfrak{R}$$

ergibt sich zunächst die Permutationsdarstellung:

$$\xi \rightarrow (1\ 2)(3\ 4), \quad \delta \rightarrow (1\ 2)(3\ 4), \quad \eta \rightarrow (1\ 3)(2\ 4).$$

Unter den möglichen Darstellungen von \mathfrak{R} wählen wir die zweidimensionale als interessanteste:

$$D(\xi^2) = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}, \quad D(\xi^3\delta) = \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}.$$

Dann kommt nach (18) und (23):

$$D^1(\xi) = D(\xi\varepsilon) = D(\varepsilon) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad D^1(\delta) = D(\xi\xi^3\delta) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad D^1(\eta) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$D^2(\xi) = D(\xi\xi) = D(\xi^2) = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}, \quad D^2(\delta) = D(\xi^3\delta) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad D^2(\eta) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$D^3(\xi) = D(\xi\eta) = D(\xi\eta\varepsilon) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad D^3(\delta) = D(\xi\eta\xi^3\delta) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad D^3(\eta) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$D^4(\xi) = D(\xi\xi\eta) = D(\eta\xi^2) = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}, \quad D^4(\delta) = D(\eta\xi^3\delta) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad D^4(\eta) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Schließlich sind noch die λ_r zu bestimmen nach der Formel

$$\lambda_r(t) = e^{2\pi i(q_r \frac{a_r}{t} + 0)}.$$

Es ergibt sich für $t = (n_1, n_2, n_3)$

$$\lambda_1 = \lambda_1(t) = e^{2\pi i (\frac{2}{3}n_1 + \frac{1}{3}n_2 + zn_3)}$$

$$\lambda_2 = \lambda_2(t) = e^{2\pi i (\frac{1}{3}n_1 + \frac{2}{3}n_2 + zn_3)}$$

$$\lambda_3 = \lambda_3(t) = e^{2\pi i (\frac{2}{3}n_1 + \frac{1}{3}n_2 - zn_3)}$$

$$\lambda_4 = \lambda_4(t) = e^{2\pi i (\frac{1}{3}n_1 + \frac{2}{3}n_2 - zn_3)}.$$

Dabei ist φ innerer Punkt von C , also $0 < z < \frac{1}{2}$. Die Darstellung von \mathfrak{G} hat also folgende Gestalt:

$$\begin{array}{l}
 t\xi \rightarrow \left[\begin{array}{c} \boxed{\begin{array}{cc} 0 & -\lambda_1 \\ \lambda_1 & -\lambda_1 \end{array}} \\ \boxed{\begin{array}{cc} \lambda_2 & 0 \\ 0 & \lambda_2 \end{array}} \end{array} \right] \quad t\eta \rightarrow \left[\begin{array}{c} \boxed{\begin{array}{cc} \lambda_1 & 0 \\ 0 & \lambda_1 \end{array}} \\ \boxed{\begin{array}{cc} \lambda_3 & 0 \\ 0 & \lambda_3 \end{array}} \\ \boxed{\begin{array}{cc} \lambda_4 & 0 \\ 0 & \lambda_4 \end{array}} \end{array} \right] \\
 t\delta \rightarrow \left[\begin{array}{c} \boxed{\begin{array}{cc} 0 & \lambda_1 \\ \lambda_1 & 0 \end{array}} \\ \boxed{\begin{array}{cc} 0 & \lambda_2 \\ \lambda_2 & 0 \end{array}} \\ \boxed{\begin{array}{cc} 0 & \lambda_3 \\ \lambda_3 & 0 \end{array}} \\ \boxed{\begin{array}{cc} 0 & \lambda_4 \\ \lambda_4 & 0 \end{array}} \end{array} \right] \quad t \rightarrow \left[\begin{array}{c} \boxed{\begin{array}{cc} \lambda_1 & 0 \\ 0 & \lambda_1 \end{array}} \\ \boxed{\begin{array}{cc} \lambda_2 & 0 \\ 0 & \lambda_2 \end{array}} \\ \boxed{\begin{array}{cc} \lambda_3 & 0 \\ 0 & \lambda_3 \end{array}} \\ \boxed{\begin{array}{cc} \lambda_4 & 0 \\ 0 & \lambda_4 \end{array}} \end{array} \right]
 \end{array}$$

Literatur.

- [1] B. L. van der Waerden, Gruppen von linearen Transformationen. Berlin 1935.
- [2] J. J. Burckhardt, Zur Theorie der Bewegungsgruppen. Comment. math. helv. 6 (1933).
- [3] F. Seitz, On the Reduction of Space groups. Ann. of Math. (2) 37 (1936).
- [4] F. Seitz, A matrix-algebraic developpement of the crystallographic groups. Zeitschr. f. Krist. 88 (1934); 90 (1935); 91 (1935); 94 (1936).
- [5] Internationale Tabellen zur Bestimmung von Kristallstrukturen. 1. Bd. Berlin, Gebr. Borntraeger, 1935.
- [6] P. Niggli u. W. Nowacki, Der arithmetische Begriff der Kristallklasse usw. Zeitschr. f. Krist. 91 (1935).

(Eingegangen am 14. 3. 1941.)

Fixpunktklassen.

Teil II.

Homotopieinvarianten der Fixpunkttheorie.

Von

Franz Wecken in Marburg a. d. Lahn.

Einleitung.

Die Ergebnisse der ersten Mitteilung¹⁾ werden in diesem Teil verschärft und erweitert, wobei besonders die Tendenz leitend ist, möglichst weitgehende homotopieinvariante Aussagen über eine Abbildung f zu erhalten²⁾.

Die Fixpunktzahlen³⁾ v_k der Fixpunktklassen wurden in W. I lediglich als eine ungeordnete, endliche Menge ganzer Zahlen gewonnen. Zur Präzisierung dieses Ergebnisses erscheint es notwendig, diese Zahlen als die Fixpunktzahlen ganz bestimmter Fixpunktklassen in einer invarianten Weise individuell zu benennen, indem man die Fixpunktklassen deformationsinvarianten Größen eindeutig zuordnet. Dies gelingt nach Nielsen mit Hilfe der Fundamentalgruppe Γ von \mathfrak{P} und eines durch f induzierten Automorphismus H von Γ .

Durch die stetige Abbildung f von \mathfrak{P} in sich wird eine Transformation von Γ in sich bekanntlich⁴⁾ nur bis auf einen inneren Automorphismus festgelegt; einen bestimmten Automorphismus H erhält man, sobald man einen Punkt p von \mathfrak{P} durch eine Kurve \mathfrak{C} mit seinem Bildpunkt $f(p) = p'$ verbindet und dadurch die Wegegruppen mit den Aufpunkten p und p' eindeutig aufeinander bezieht. H ist mit dem System $\{f, \mathfrak{C}\}$ deformationsinvariant

¹⁾ Fixpunktklassen. I. Math. Annalen 117 (1941), S. 659–671, zit. als W. I. Dort findet sich in § 1 die Erklärung der häufig benutzten Bezeichnungen. Druckfehlerberichtigung zu W. I: auf S. 663, Zeile 13 v. o. lies $\overline{\mathfrak{H}} - \overline{\mathfrak{H}} \cap \mathfrak{H}$ statt $\overline{\mathfrak{H}} - \overline{\mathfrak{H}} \cap \mathfrak{H}$; Zeile 16 v. o. lies „obere“ statt „untere“. — Weitere Literatur: P. Alexandroff u. H. Hopf, Topologie I. Berlin 1935 (zit. als AH.); J. Nielsen, Untersuchungen zur Topologie der geschlossenen zweiseitigen Flächen I. Acta math. 50 (1927), S. 189–358 (zit. als N.); K. Reidemeister, Automorphismen von Homotopiekettenringen. Math. Annalen 112 (1936), S. 586–593 (zit. als R. 1); K. Reidemeister, Topologie der Polyeder und kombinatorische Topologie der Komplexe. Leipzig 1938 (zit. als R. 2); A. Speiser, Die Theorie der Gruppen von endlicher Ordnung, 3. Aufl. Berlin 1937 (zit. als S.); H. Seifert und W. Threlfall, Lehrbuch der Topologie. Leipzig 1934 (zit. als ST.).

²⁾ Vgl. den Vorbericht in W. I, S. 661.

³⁾ Der Ausdruck „Fixpunktzahl“ bedeutet in dieser Mitteilung stets die algebraische Fixpunktzahl.

⁴⁾ Vgl. ST., S. 155f., 176.

verknüpft oder ist, wie man auch sagen kann, auf der Abbildungsklasse f im kleinen invariant. Die Gesamtheit der so durch f erzeugbaren Automorphismen, die sogenannte *Automorphismenfamilie* Σ_1 , ist dagegen eine homotopieinvariante Eigenschaft von f oder eine Invariante von f im großen (§ 1). Es erweist sich weiterhin als zweckmäßig, im kleinen invariante Eigenschaften zu untersuchen und hieraus später die angestrebten Invarianten im großen abzuleiten.

Die durch H in I' erzeugten H -Klassen p_k^H entsprechen bei geeigneter Definition eineindeutig den sämtlichen Fixpunktklassen bei f , und dieser Sachverhalt bleibt bei Deformation von $\{f, \mathbb{C}\}$ bestehen (§ 2). Durch Kombination der p_k^H mit den zugehörigen Fixpunktzahlen v_k wird so ein deformationsinvarianter Ausdruck R_0 erhalten, der mehr als die Gesamtheit der v_k beinhaltet.

Für die folgenden Untersuchungen wird als neues Hilfsmittel die *universelle Überlagerung* $\tilde{\mathbb{P}}$ von \mathbb{P} eingeführt (§ 3), auf die die ursprüngliche Nielsensche Definition der Fixpunktklassen (N., S. 289) wesentlich Bezug nimmt, die jedoch in § 1 und 2 sowie in Teil I und III nicht benutzt wird. Hierdurch wird ein tieferes Verständnis des Vorangegangenen vermittelt; z. B. wird erkennbar, daß die Deformation des Systems $\{f, \mathbb{C}\}$ nichts anderes bedeutet als die gleichzeitige Deformation von $f(x)$ und einer darüberliegenden stetigen Abbildung $F(X)$ von $\tilde{\mathbb{P}}$ in sich; eine Fixpunktklasse bei $f(x)$ erscheint als die Punktmenge, über der die bei einer Abbildung $F(X)$ auftretenden Fixpunkte liegen. Gleichzeitig wird durch die Einführung von $\tilde{\mathbb{P}}$ die Brücke geschlagen zu einer Methode der kombinatorischen Topologie, die an den Begriff des universellen Überlagerungskomplexes anknüpft und zur Zeit nicht durch äquivalente stetigkeitstopologische Verfahren ersetzbar ist, nämlich zur Theorie des *Homotopiekettenringes* eines Komplexes K .

Eine simpliziale Abbildung von \mathbb{P} in sich bestimmt eine Transformation der gewöhnlichen Kettengruppen eines Komplexes K in sich. Eine darüberliegende Abbildung des universellen Überlagerungskomplexes \tilde{K} in sich transformiert die Ketten von \tilde{K} oder Homotopieketten von K in sich, und zwar vermöge ihrer Herkunft in spezieller Weise. Es entsteht so ein *Automorphismus* T des Homotopiekettenringes (Reidemeister, R. 1), der also das kombinatorische Äquivalent einer bestimmten über f liegenden Selbstabbildung von \mathbb{P} (entsprechend einem System $\{f, \mathbb{C}\}$) darstellt.

Die aus den Matrizen von T von Reidemeister gewonnene *Spureninvariante* R wird nun in § 4 als mit der für die simpliziale Abbildung gebildeten Größe R_0 *identisch* nachgewiesen. Es ist damit gezeigt, wie sich R_0 aus einer geeigneten Simplizialapproximation von f berechnen läßt.

Aus R_0 lassen sich alle Fixpunktzahlen v_k , daher auch die Lefschetzsche Zahl $A = \sum_k v_k$ und die Zahl μ der wesentlichen Fixpunktklassen ablesen. Hierdurch ist eine *verbesserte untere Schranke* für die *Fixpunktmindestzahl* m gewonnen, denn nach W. 1 gilt ja $m \geq \mu$, während aus der Kenntnis von A lediglich im Falle $A \neq 0$ die Abschätzung $m \geq 1$ folgt. Wie in Teil III gezeigt werden wird, gilt unter den dortigen Voraussetzungen $m = \mu$; die aus der Homotopietheorie hergeleitete Invarianzaussage verschärft also das Ergebnis der Homologietheorie wesentlich, indem sie die *genaue* untere Schranke der geometrischen Fixpunktzahl liefert.

In § 5 wird durch Berücksichtigung des Einflusses der speziell gewählten Kurve \mathbb{C} auf H und R_0 ein im großen homotopieinvarianter Ausdruck Σ_2 gewonnen, der die bisher erwähnten Invarianten umfaßt.

Schon nach der Definition der durch einen Deformationsweg geleisteten Zuordnung zwischen den Fixpunktklassen homotoper Abbildungen (W. 1, S. 669) lag die Frage nahe, ob und unter welchen Bedingungen diese Zuordnung vom Wege unabhängig sei; doch erfordert ein genaueres Eingehen auf sie die in den vorigen Paragraphen entwickelten formalen Hilfsmittel. Ein Beispiel in § 6 zeigt, daß die Zuordnung vom Deformationswege wesentlich abhängen kann; es können also durch geschlossene Wege auf f Permutationen unter den Fixpunktklassen bei f erzeugt werden. Dafür, daß dies nicht eintritt, wird eine hinreichende Bedingung angegeben, die in wichtigen Fällen erfüllt ist.

Bei den erwähnten geschlossenen Deformationswegen $\{f_\tau\}$ beschreibt jeder Bildpunkt $f_\tau(p)$ eines festen Punktes p ebenfalls geschlossene Wege; die durch sie bestimmte Untergruppe g_3 von Γ , die deformationsinvariant mit $\{f, \mathbb{C}\}$ verknüpft ist, bestimmt die auftretenden Permutationen (§ 7). Aus Σ_2 wird durch Hinzunahme von g_3 eine erweiterte Invariante Σ_3 gewonnen.

In § 8 werden zur Veranschaulichung der Begriffe die Invarianten für die Abbildungen des Kreises in sich berechnet.

§ 1.

Die durch Abbildungen von \mathfrak{P} erzeugten Automorphismen von Γ .

Es sei p ein Punkt des Polyeders \mathfrak{P} . Die in p beginnenden geschlossenen Kurven \mathfrak{W} — homotope Kurven als nicht verschieden gerechnet — bilden bekanntlich⁵⁾ die *Wegegruppe* oder *Fundamentalgruppe* Γ von \mathfrak{P} . Die Klasse der zu \mathfrak{W} homotopen Wege bezeichnen wir mit $[\mathfrak{W}]$, im übrigen die Elemente von Γ mit α, β, \dots ⁶⁾. Nicht notwendig geschlossene Kurven werden mit \mathbb{C} bezeichnet. Auch für sie ist (bei Festhaltung der Endpunkte, vgl. W. 1,

⁵⁾ Zur Theorie der Fundamentalgruppe vgl. ST., Kap. VII.

⁶⁾ Geschweifte Klammern $\{\}$, die bei ST. Wegeklassen bezeichnen, benutzen wir für Mengen allgemeiner Art.

S. 665) eine Homotopie und die Bildung von Klassen $[\mathbb{C}]$ erklärt. Für das formale Rechnen mit diesen Klassen bestehen einfache Regeln⁷⁾, die wir im folgenden stillschweigend benutzen. Die Klasse der in einem Punkte q beginnenden geschlossenen und zusammenziehbaren Wege bezeichnen wir mit $[q]$.

$f(x) = x'$ sei eine stetige Selbstabbildung von \mathfrak{P} und Γ_p die Wegegruppe von \mathfrak{P} mit dem Aufpunkt p . Nach den Überlegungen in ST. (Schluß von § 42) wird durch f eine homomorphe Abbildung⁸⁾ von Γ_p in $\Gamma_{p'}$ erzeugt. Da Γ_p und $\Gamma_{p'}$ isomorph sind, kann man von einem Autohomomorphismus (wir sagen kurz *Automorphismus*) H von Γ sprechen, sobald Γ_p und $\Gamma_{p'}$ durch eine von p nach p' verlaufende Kurve \mathbb{C} eindeutig aufeinander bezogen sind. Als Automorphismus von Γ_p läßt sich H definieren durch

$$(1) \quad H[\mathfrak{B}] = [\mathbb{C} f(\mathfrak{B}) \mathbb{C}^{-1}],$$

wenn \mathfrak{B} ein geschlossener in p beginnender Weg und $f(\mathfrak{B})$ sein Bild ist. H ist also durch f und \mathbb{C} eindeutig festgelegt und ändert sich auch nicht (ST., § 50), wenn wir f , \mathbb{C} und p einzeln oder gleichzeitig deformieren, wenn wir nur dafür sorgen, daß stets zwischen p und seinem Bildpunkt Verbindung gehalten wird. Wir sagen also, daß H dem System $\{f, \mathbb{C}\}$ *deformationsinvariant* zugeordnet ist. Allgemeiner kann man sagen: Wenn der Abbildung f (durch zusätzliche Angaben) ein bestimmter Automorphismus H von Γ zugeordnet ist, so läßt sich diese Zuordnungsvorschrift stets auf benachbarte Abbildungen g ausdehnen, und der Automorphismus ist daher auf der Abbildungsklasse \mathfrak{f} *im kleinen invariant*.

Schreibt man \mathbb{C} nicht vor, so ist die Transformation von Γ nur bis auf einen inneren Automorphismus T_γ bestimmt (ST., § 42); genauer: ersetzt man \mathbb{C} durch \mathbb{C}_2 mit $[\mathbb{C}_2] = \gamma[\mathbb{C}]$, so erhält man statt (1)

$$(2) \quad [\mathbb{C}_2 f(\mathfrak{B}) \mathbb{C}_2^{-1}] = \gamma(H[\mathfrak{B}])\gamma^{-1} = T_\gamma H[\mathfrak{B}]$$

mit $T_\gamma \alpha = \gamma \alpha \gamma^{-1}$, H ist also durch $T_\gamma H$ ersetzt. Die Gesamtheit

$$\Sigma_1 = \{T_\gamma H \mid \gamma \in \Gamma\}$$

der so erhältlichen Automorphismen nennen wir eine *Automorphismenfamilie*⁹⁾. Jeder Automorphismus aus Σ_1 kann durch jede Abbildung aus \mathfrak{f} erzeugt werden, Σ_1 ist demnach *auf \mathfrak{f} im großen invariant* oder ist eine *Homotopie-invariante* der Abbildungen.

⁷⁾ Die $[\mathbb{C}]$ bilden ein *Gruppoid* nach H. Brandt. Über eine Verallgemeinerung des Gruppenbegriffes. Math. Annalen 96 (1927), S. 360–366; vgl. auch S., S. 4ff.

⁸⁾ Vgl. S., § 9: — Statt „mehrstufig isomorph“ sagen wir stets „homomorph“.

⁹⁾ Dieser Wortsinn ist weiter als bei Nielsen (N., S. 244), der nur Autoisomorphismen betrachtet.

Bei spezieller Gestalt von \mathfrak{P} erweist sich Σ_1 als so gehaltvolle Eigenschaft von \mathfrak{f} , daß \mathfrak{f} durch Σ_1 schon völlig festgelegt wird¹⁰⁾ und alle weiteren Aussagen über \mathfrak{f} aus Σ_1 folgen; allgemein ist dies jedoch nicht der Fall.

§ 2.

Zuordnung zwischen Fixpunktklassen und H -Klassen.

Haben p , f und \mathfrak{C} die obige Bedeutung, so kann man jedem Fixpunkt q bei f gewisse Elemente $\alpha \in \Gamma$ zuordnen, indem man eine Kurve \mathfrak{C}_1 von p nach q wählt und

$$(3) \quad \alpha = [\mathfrak{C} f(\mathfrak{C}_1) \mathfrak{C}_1^{-1}]$$

bildet (Fig. 1). Jedem Fixpunkt entspricht also eine nicht leere Menge von Gruppenelementen α , und jedem $\alpha \in \Gamma$ eine (vielleicht leere) Menge von Fixpunkten.

Bezeichnet man nach Reidemeister (R. 1, S. 589) zwei Elemente $\alpha, \beta \in \Gamma$ als konjugiert bezüglich H , wenn

$$(4) \quad \beta = (H\gamma)\alpha\gamma^{-1}$$

mit passendem $\gamma \in \Gamma$ ist, so erhält man eine Einteilung von Γ in Klassen konjugierter Elemente, kurz H -Klassen¹¹⁾.

Hilfssatz. Die angegebene Vorschrift (3) ordnet jedem Fixpunkt genau eine ganze H -Klasse, jedem Gruppenelement entweder nichts oder eine ganze Fixpunktklasse zu. Diese Zuordnung bleibt bei Deformation von $\{f, \mathfrak{C}\}$ bestehen, soweit die Fixpunktklassen erhalten bleiben. Durch Deformation läßt sich erreichen, daß einem vorgeschriebenen Gruppenelement eine Fixpunktklasse zugeordnet wird.

Beweis. Wir ersetzen in (3) \mathfrak{C}_1 durch einen beliebigen Weg von p nach q . Jeder solche Weg kann durch homotope Deformation, die hier offenbar belanglos ist, in die Form $\mathfrak{W}\mathfrak{C}_1$ gebracht werden, und $[\mathfrak{W}] = \gamma$ kann hierbei jedes Element aus Γ sein. Mit $\mathfrak{W}\mathfrak{C}_1$ statt \mathfrak{C}_1 erhält man nun nach (3) und (1) statt α ein Element

$$\begin{aligned} \beta &= [\mathfrak{C} f(\mathfrak{W}) f(\mathfrak{C}_1) \mathfrak{C}_1^{-1} \mathfrak{W}^{-1}] = [\mathfrak{C} f(\mathfrak{W}) \mathfrak{C}^{-1}] [\mathfrak{C} f(\mathfrak{C}_1) \mathfrak{C}_1^{-1}] [\mathfrak{W}^{-1}] \\ &= (H[\mathfrak{W}])\alpha[\mathfrak{W}^{-1}] = (H\gamma)\alpha\gamma^{-1}. \end{aligned}$$

Dieser Ausdruck durchläuft in der Tat nach (4) eine H -Klasse. — Sind q_1 und q_2 Fixpunkte von gleicher Fixpunktklasse, so sei die Kurve \mathfrak{C}_0 von q_1

¹⁰⁾ N., §§ 22, 27. Nielsen benutzt diesen Umstand zur Definition der Abbildungsklasse.

¹¹⁾ Nach Nielsen (N., S. 260) heißen in diesem Falle α^{-1} und β^{-1} isogredient; an der Stelle der H -Klassen stehen dort die Isogredientenklassen.

nach q_2 gewählt mit $[\mathfrak{C}_0] = [f(\mathfrak{C}_0)]$. Man erhält dann nach (3) für q_1 und q_2 dasselbe Element α , sobald man von p nach q irgendeinen Weg \mathfrak{C}_1 und von p nach q_2 den Weg $\mathfrak{C}_2 = \mathfrak{C}_1 \mathfrak{C}_0$ wählt (Fig. 2). Sind umgekehrt q_1 und q_2 zwei Fixpunkte bei f , denen dasselbe Element

$$[\mathfrak{C}/(\mathfrak{C}_1 \mathfrak{C}_1^{-1})] = [\mathfrak{C}/(\mathfrak{C}_2 \mathfrak{C}_2^{-1})]$$

zugeordnet ist, so folgt sofort

$$[f(\mathfrak{C}_1) \mathfrak{C}_1^{-1}] = [f(\mathfrak{C}_2) \mathfrak{C}_2^{-1}],$$

$$[\mathfrak{C}_1^{-1} \mathfrak{C}_2] = [f(\mathfrak{C}_1^{-1} \mathfrak{C}_2)],$$

und q_1 und q_2 müssen der gleichen Fixpunktclass angehören, da der Weg $\mathfrak{C}_1^{-1} \mathfrak{C}_2$ von q_1 nach q_2 führt (Fig. 3).

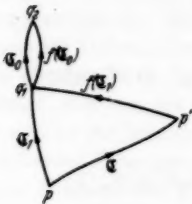


Fig. 2.

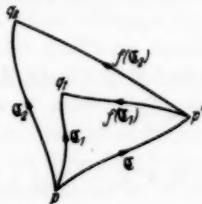


Fig. 3.

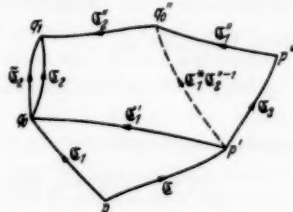


Fig. 4.

Sind durch einen Deformationsweg zwei Fixpunktclassen verschiedener Abbildungen einander zugeordnet (W. I, § 3), so ist zu zeigen, daß ihnen dieselbe H -Klasse zukommt. Es sei also $\{f_\tau \ (0 \leq \tau \leq 1)\}$ ein Deformationsweg von $f = f_0$ nach f_1 , es sei weiter $f_0(q_0) = q_0$, $f_1(q_1) = q_1$, $\mathfrak{C}_1 = \{q_\tau \ (0 \leq \tau \leq 1)\}$, $\mathfrak{C}_2 = \{f_\tau(q_\tau) \ (0 \leq \tau \leq 1)\}$,

$$(5) \quad \{\tilde{\mathfrak{C}}_2\} = \{\mathfrak{C}_2\}.$$

Wir schreiben $f_0(x) = x'$, $f_1(x) = x''$, entsprechend sei $\mathfrak{C}'_1, \mathfrak{C}''_1$ erklärt (Fig. 4). $\mathfrak{C}_1 = \{p_\tau \ (0 \leq \tau \leq 1)\}$ sei ein Weg von p nach q_0 , $\mathfrak{C}_3 = \{f_\tau(p) \ (0 \leq \tau \leq 1)\}$. Als Weg von p nach q_1 wählen wir $\mathfrak{C}_1 \mathfrak{C}_2$ und haben nun zu zeigen, daß die beiden Wegeklassen

$$(6) \quad [\mathfrak{C} \mathfrak{C}'_1 \mathfrak{C}_1^{-1}], \quad [\mathfrak{C} \mathfrak{C}_3 \mathfrak{C}'_1 \mathfrak{C}_2 \mathfrak{C}_2^{-1} \mathfrak{C}_1^{-1}]$$

bezüglich H konjugiert sind; wir werden sogar finden, daß beide Klassen identisch sind.

Durch die stetige Kurvenschar

$$\mathfrak{C}_\sigma^* = \{f_{\sigma\tau}(p_\tau) \ (0 \leq \tau \leq 1); \ f_{\sigma+\tau-\sigma\tau}(q_\tau) \ (0 \leq \tau \leq 1)\} \ (0 \leq \sigma \leq 1)$$

wird $\mathbb{C}_0^* = \mathbb{C}_1^* \tilde{\mathbb{C}}_2$ in

$$\mathbb{C}_1^* = \{f_\tau(p_\tau) \ (0 \leq \tau \leq 1); \ f_1(q_\tau) \ (0 \leq \tau \leq 1)\}$$

deformiert, es ist also nach (5)

$$(7) \quad [\mathbb{C}_1^* \mathbb{C}_2] = [\mathbb{C}_1^* \tilde{\mathbb{C}}_2] = [\mathbb{C}_1^*].$$

Weiter wird die Kurve \mathbb{C}_1^* , die sich auch als

$$\mathbb{C}_1^* = \left\{ f_{\frac{\tau}{2}}(p_{\frac{\tau}{2}}) \ (0 \leq \tau \leq 1); \ f_{\frac{1+\tau}{2}}(p_{\frac{1+\tau}{2}}) \ (0 \leq \tau \leq 1); \ f_1(q_\tau) \ (0 \leq \tau \leq 1) \right\}$$

schreiben läßt, durch

$$\mathbb{C}_\sigma^* = \left\{ f_{\frac{\sigma}{2}}(p_{\frac{2-\sigma}{2}}) \ (0 \leq \tau \leq 1); \ f_{1-(1-\sigma)\frac{2-\sigma}{2}}(p_{1-\frac{1-\tau}{2}\sigma}) \ (0 \leq \tau \leq 1); \right. \\ \left. f_1(q_\tau) \ (0 \leq \tau \leq 1) \right\} \quad (1 \leq \sigma \leq 2)$$

stetig überführt in

$$\mathbb{C}_2^* = \{f_\tau(p_0) \ (0 \leq \tau \leq 1); \ f_1(p_\tau) \ (0 \leq \tau \leq 1); \ f_1(q_\tau) \ (0 \leq \tau \leq 1)\} \\ = \mathbb{C}_3 \mathbb{C}_1'' \mathbb{C}_2'',$$

also folgt nach (7)

$$[\mathbb{C}_1^* \mathbb{C}_2] = [\mathbb{C}_1^*] = [\mathbb{C}_2^*] = [\mathbb{C}_3 \mathbb{C}_1'' \mathbb{C}_2''], \quad [\mathbb{C}_1^*] = [\mathbb{C}_3 \mathbb{C}_1'' \mathbb{C}_2''^{-1}],$$

und damit die Identität der Ausdrücke (6). —

Zum Beweise der dritten Behauptung nehmen wir $\{f, \mathbb{C}\}$ sowie ein Element $\gamma \in \Gamma$ als gegeben an und haben nun einen solchen bei $f = f_0$ beginnenden Deformationsweg $\{f_\tau \ (0 \leq \tau \leq 1)\}$ anzugeben, daß einem bei f_1 auftretenden Fixpunkt das Element γ zugeordnet wird. $\mathbb{C}_0 = \{p_\tau \ (0 \leq \tau \leq 1)\}$ sei ein Weg von $f(p)$ nach p , so daß $[\mathbb{C} \mathbb{C}_0] = \gamma$ ist. Den Deformationsweg wählen wir so, daß

$$(8) \quad f_\tau(p) = p_\tau \ (0 \leq \tau \leq 1)$$

ist; die Existenz werden wir sogleich zeigen. p ist dann Fixpunkt bei f_1 . Als Weg von p nach $f_1(p)$ (\mathbb{C} in Formel (3)) haben wir $\mathbb{C} \mathbb{C}_0$ zu verwenden; den Weg \mathbb{C}_1 in (3) unterdrücken wir, da der Fixpunkt mit p zusammenfällt. Dann wird nach (3) in der Tat $\alpha = [\mathbb{C} \mathbb{C}_0] = \gamma$. — Um die Existenz von $\{f_\tau\}$ einzusehen, führen wir in einer geeigneten Umgebung von p „Polarkoordinaten“ r, φ ein¹²⁾; $u(r, \varphi)$ sei für $0 \leq r \leq 2\varepsilon$ der Punkt mit den Polarkoordinaten r

¹²⁾ Es sei eine Zerlegung K von \mathfrak{P} gewählt, bei der p als Ecke auftritt; der offene Simplexstern $O_K(p)$ (vgl. AH., S. 131) enthalte alle $x \in \mathfrak{P}$ mit $\varrho(x, p) \leq 2\varepsilon$. Setzt man den Abstand $\varrho(x, p) = r(x)$ und versteht unter $\varphi(x)$ für $r(x) \leq 2\varepsilon$ den Punkt des Begrenzungsakomplexes $B_K(p)$ (nach AH., S. 131; des „Ringes um p “ nach R. 2, S. 50), in den x von p aus projiziert wird, so sind $r(x)$ und $\varphi(x)$ als Polarkoordinaten von x verwendbar.

und φ ; er durchfegt die 2ε -Umgebung von p . Dann sei für $0 \leq \tau \leq 1$ (mit $r = r(x)$, $\varphi = \varphi(x)$)

$$f_\tau(x) = \begin{cases} p_{\tau - \frac{r}{\varepsilon}} & \text{für } 0 \leq \tau \leq \varepsilon\tau, \\ f(u(r - \varepsilon\tau, \varphi)) & \text{für } \varepsilon\tau \leq \tau \leq \varepsilon, \\ f(u(\tau + r - 2\varepsilon\tau, \varphi)) & \text{für } \varepsilon \leq \tau \leq 2\varepsilon, \\ f(x) & \text{sonst für } x \in \mathfrak{P}. \end{cases}$$

Die Stetigkeit in τ und x sowie die Bedingungen (8) und $f_0 = f$ sind leicht zu verifizieren. Damit ist der Hilfssatz bewiesen.

Da die Zahl der Fixpunktklassen stets endlich ist (W. 1, Satz 1), die Zahl der H -Klassen p_1^H, p_2^H, \dots jedoch unendlich sein kann (hierfür sind sehr leicht Beispiele zu bilden), gibt es im allgemeinen H -Klassen ohne zugeordnete Fixpunktklassen. Eine einer H -Klasse p_k^H zugeordnete Fixpunktklasse f_k kann bei Deformation der Abbildung verschwinden und wieder auftauchen; da sie aber durch die H -Klasse p_k^H individuell kenntlich ist, ist es gerechtfertigt, stets von der zu p_k^H gehörigen Fixpunktklasse f_k zu sprechen, indem man f_k als leer bezeichnet, wenn keine Fixpunkte p_k^H zugeordnet sind. Zuzufolge der letzten Behauptung des Hilfssatzes kommt dann bei jeder Abbildung aus f jeder H -Klasse eine Fixpunktklasse zu. Jede leere Fixpunktklasse soll durch Definition die Fixpunktzahl 0 erhalten. Diese erweiterte Definition der Fixpunktklassen ermöglicht es, die Sätze 2 und 3 aus W. 1 in verschärfter Form auszusprechen, und wir gewinnen zusammenfassend folgendes Ergebnis:

Satz 1. *Nach Wahl von f und \mathfrak{C} besteht zufolge (3) eine eindeutige und deformationsinvariante Zuordnung zwischen den H -Klassen und den Fixpunktklassen. Die algebraischen Fixpunktzahlen aller Fixpunktklassen sind deformationsinvariant. Bei jeder Abbildung sind alle Fixpunktklassen bis auf endlich viele leer, aber keine Fixpunktklasse bleibt bei Deformation der Abbildung beständig leer.*

Nach Satz 1 erscheinen die Fixpunktzahlen v_k mittelbar den H -Klassen p_k^H zugeordnet. Man kann in vektorieller Schreibweise¹³⁾

$$(9) \quad R_0 = \sum_k v_k p_k^H$$

die H -Klassen mit den zugeordneten Fixpunktzahlen zusammenfassen. In der Kombination

$$(10) \quad \{H, R_0\}$$

sind alle von uns abgeleiteten deformationsinvarianten Eigenschaften von $\{f, \mathfrak{C}\}$ enthalten.

¹³⁾ Die additive Auffassung wird später (§ 4) gerechtfertigt werden.

§ 3.

Eine andere Definition der Fixpunktklassen.

Die von Nielsen bevorzugte geometrische Deutung der (von ihm rein algebraisch eingeführten) Fixpunktklassen ist nicht die von uns angegebene. Wir benötigen sie jetzt, um die Identität von R_0 mit der Reidemeisterschen Spureninvarianten nachzuweisen (§ 4), und führen dazu die *universelle Überlagerung* $\tilde{\mathfrak{P}}$ von \mathfrak{P} ein.

Wir definieren in üblicher Weise¹⁴⁾ $\tilde{\mathfrak{P}}$ als die *Gesamtheit der Klassen in p beginnender Wege in \mathfrak{P}* , wo p wie bisher ein fest gewählter Punkt von \mathfrak{P} ist¹⁵⁾. Als *Punkte* von $\tilde{\mathfrak{P}}$ bezeichnen wir jene Wegeklassen auch mit P, Q, X, Y, \dots ; insbesondere sei $[p] = P$. Die Abhängigkeit der Überlagerung $\tilde{\mathfrak{P}} = \tilde{\mathfrak{P}}_p$ von p ist nur eine scheinbare, da die Klassen von $\tilde{\mathfrak{P}}_q$ aus denen von $\tilde{\mathfrak{P}}_p$ gewonnen werden können durch eine triviale Linksmultiplikation, die alle wichtigen Eigenschaften von $\tilde{\mathfrak{P}}$ erhält.

Es sei $q \in \mathfrak{P}$, \mathfrak{C}_0 ein Weg von p nach q ; wir setzen $[\mathfrak{C}_0] = Q$. Jeder solche Punkt von $\tilde{\mathfrak{P}}$ erlaubt eine Multiplikation von links mit Elementen aus Γ :

$$\alpha Q = [\mathfrak{A}\mathfrak{C}_0],$$

wenn $\alpha = [\mathfrak{A}]$ ist. Γ ist somit gleichzeitig eine Gruppe von Selbstabbildungen (Decktransformationen) von $\tilde{\mathfrak{P}}$. Die Punkte αQ mit $\alpha \in \Gamma$ heißen die *über q liegenden Punkte* von $\tilde{\mathfrak{P}}$. Topologische und metrische Eigenschaften von \mathfrak{P} lassen sich auf $\tilde{\mathfrak{P}}$ übertragen, z. B. durch die Festsetzung, daß als *Entfernung* $\bar{\rho}(Q_1, Q_2)$ zweier Punkte von $\tilde{\mathfrak{P}}$ die untere Grenze der Durchmesser aller Wege \mathfrak{C} mit $Q_1[\mathfrak{C}] = Q_2$ gelten soll. $\tilde{\mathfrak{P}}$ wird hierdurch zu einem metrischen Raum. Die Decktransformationen sind stetig, isometrisch und fixpunktfrei.

Durch die Funktion

$$\Phi(Q) = q,$$

die jedem Punkt $Q = [\mathfrak{C}]$ den Endpunkt q der Wege \mathfrak{C} zuordnet, wird $\tilde{\mathfrak{P}}$ stetig und im kleinen topologisch auf \mathfrak{P} abgebildet. Durch ihre Umkehrung Φ^{-1} wird jedem Wege \mathfrak{C} in \mathfrak{P} eindeutig ein über \mathfrak{C} liegender Weg $\tilde{\mathfrak{C}}$ in $\tilde{\mathfrak{P}}$ zugeordnet, sobald dessen Anfangspunkt aus den über dem Anfangspunkt von \mathfrak{C} liegenden Punkten gewählt ist. Hierdurch wird es auch möglich, jede Simplicialzerlegung von \mathfrak{P} nach $\tilde{\mathfrak{P}}$ auszudrücken; $\tilde{\mathfrak{P}}$ ist, wie man alsbald erkennt, homöomorph

¹⁴⁾ Vgl. H. Weyl, Die Idee der Riemannschen Fläche. Leipzig 1913. S. 51. B. v. Kerékjártó, Vorlesungen über Topologie I. Berlin 1923. S. 176.

¹⁵⁾ Daß die Klassen geschlossener Wege aus p gleichzeitig Gruppenelemente sind, wird nicht zu Mißverständnissen Anlaß geben.

zu einem lokal-endlichen Polyeder¹⁶⁾ von gleicher Dimensionszahl wie \mathfrak{P} . ist also ein „Komplex“¹⁷⁾. Ein einbettender euklidischer Raum wird durch diese Konstruktion nicht geliefert.

Ist eine stetige Abbildung $f(x)$ von \mathfrak{P} in sich gegeben und ein Punkt $P' = [\mathfrak{C}]$ über $f(p)$ gewählt, so gibt es genau eine stetige Selbstabbildung $F(X)$ von $\tilde{\mathfrak{P}}$, so daß

$$(11) \quad F(P) = P'$$

ist und aus $F(X) = Y$ stets $f(\Phi(X)) = \Phi(Y)$ folgt. (Die Konstruktion von F erfolgt wieder mittels der Funktion Φ^{-1} .) Man sagt, daß $F(X)$ eine über $f(x)$ liegende Selbstabbildung von $\tilde{\mathfrak{P}}$ ist; die sämtlichen über f liegenden derartigen Abbildungen sind die Funktionen $\alpha F(X)$ mit $\alpha \in I'$. Es zeigt sich, daß mit der oben getroffenen Wahl von \mathfrak{C} gerade eine bestimmte unter den über f liegenden Abbildungen festgelegt ist.

An Hand der Überlegungen von § 1 ersieht man, daß für $\alpha \in I'$

$$F(\alpha P) = (H\alpha)F(P)$$

ist, wenn (11) vorausgesetzt ist; da $F(\alpha X)$ über $f(x)$ liegt, folgt allgemein

$$(12) \quad F(\alpha X) = (H\alpha)F(X).$$

Ist q Fixpunkt bei $f(x)$ und liegt Q über q , so muß mit passendem $\alpha \in I'$

$$(13) \quad F(Q) = \alpha Q$$

sein; α hat den in (3) angegebenen Wert, denn aus $F([\nu]) = [\mathfrak{C}]$ folgt nach Fig. 1 aus Stetigkeitsgründen

$$F([\mathfrak{C}_1]) = [\mathfrak{C}/(\mathfrak{C}_1)] = [\mathfrak{C}/(\mathfrak{C}_1)\mathfrak{C}_1^{-1}][\mathfrak{C}_1] = \alpha[\mathfrak{C}_1].$$

Die durch (13) erklärte Zuordnung zwischen Fixpunkten und Gruppenelementen ist also die uns bekannte. Zuzufolge (13) liegt über q ein Fixpunkt bei der Abbildung $\alpha^{-1}F(X)$. Ist q_1 ein Fixpunkt der gleichen Klasse wie q , ist also etwa

$$F(Q_1) = (H\gamma)\alpha\gamma^{-1}Q_1,$$

so folgt nach (12)

$$F(\gamma^{-1}Q_1) = \alpha(\gamma^{-1}Q_1).$$

d. h. auch über q_1 liegt ein Fixpunkt (nämlich $\gamma^{-1}Q_1$) bei derselben Abbildung. Daß umgekehrt jeder Fixpunkt bei $\alpha^{-1}F$ über einem Fixpunkt bei f aus der genannten Fixpunktklasse liegt, ist nach dem Gesagten klar. So ergibt sich der Satz¹⁸⁾: *Liegen über $f(x)$ die Funktionen $\alpha^{-1}F(X)$ ($\alpha \in I'$) mit den Fix-*

¹⁶⁾ Vgl. AH., S. 129 und 149.

¹⁷⁾ Im Sinne von ST., S. 42.

¹⁸⁾ N., Satz 13.

punktmengen f_α , so stellen die Punktmengen $f_\alpha = \Phi(\tilde{f}_\alpha)$ die Fixpunktclassen bei $f(x)$ dar.

Es ist $f_\alpha = f_\beta$, wenn α und β bezüglich H konjugiert sind. — Der eben ausgesprochene Satz liefert die angekündigte Definition der Fixpunktclassen, die sich rein geometrisch formulieren läßt, sobald die universelle Überlagerung $\tilde{\mathfrak{P}}$ und die über f liegenden Selbstabbildungen von $\tilde{\mathfrak{P}}$ eingeführt sind.

§ 4.

Die Reidemeistersche Spureninvariante.

Zur praktischen Berechnung der Fixpunktzahlen knüpfen wir an frühere Überlegungen an. Aus W. 1, S. 663 und A.H., S. 528, 538, 542 entnehmen wir folgenden Hilfssatz: *Es sei K_1 eine Unterteilung der Zerlegung K von \mathfrak{P} , $g(x)$ eine simpliciale Abbildung von K_1 in K mit nur Grund-Fixsimplexen und $g(c')$ der von $g(x)$ in der r -dimensionalen Kettengruppe $L'(K_1)$ induzierte Automorphismus. Jedes Fixsimplex a'_i bei $g(c')$ enthält genau einen Fixpunkt bei $g(x)$; ist j sein Index, so ist*

$$(14) \quad g(a'_i) = \sum_k l'_{ik} a'_k, \quad l'_{ii} = j.$$

Abbildungen $g(x)$ der bezeichneten Art liegen überall dicht in jeder Abbildungsklasse.

Diese Tatsachen, die in derselben Gruppierung zum Beweis des Fixpunktsatzes verwendet wurden, kombinieren wir mit dem im vorigen Paragraphen ausgesprochenen Satze.

\tilde{K}_1 , der universelle Überlagerungskomplex¹⁹⁾ von K_1 , stellt eine Zerlegung von $\tilde{\mathfrak{P}}$ dar, und eine über $g(x)$ liegende Abbildung $G(X)$ ist eine Simplicialabbildung, die eine Transformation $T(\tilde{c})$ der Ketten von \tilde{K}_1 induziert. T ist ein Automorphismus des zu K_1 gehörigen Homotopiekettenringes²⁰⁾ im Sinne von R. 1. Die Abbildung $\ast G(X)$ ($\ast \in I$) induziert den Automorphismus $\ast T(\tilde{c})$. Die Zellen von \tilde{K}_1 bezeichnen wir mit $\gamma a'_\ast$ ($\gamma \in I$) und schreiben für $1 a'_i$ wieder a'_i . Ist etwa

$$(14') \quad T(a'_i) = \sum_k l'_{ik} a'_k,$$

$g(q) = q$, $q \in a'_i$, Q über q , $\alpha^{-1}G(Q) = Q$, so ist das über a'_i liegende, Q enthaltende Simplex $\gamma^{-1}a'_i$ von \tilde{K}_1 Fixsimplex bei $\alpha^{-1}T$, und die Kette $\alpha^{-1}T(\gamma^{-1}a'_i)$ muß nach (14) das Glied

$$j\gamma^{-1}a'_i$$

¹⁹⁾ Vgl. ST., S. 194; R. 2, S. 182.

²⁰⁾ Zur Definition der Homotopieketten vgl. auch R. 2, § 17. 5.

enthalten. Wegen

$$T(\gamma \tilde{c}) = (H\gamma)T(\tilde{c})$$

[vgl. (11) und R. 1, (2)] und (14') ist aber

$$\alpha^{-1}T(\gamma^{-1}\alpha_i^r) = \alpha^{-1}(H\gamma)^{-1} \sum_k t_{ik}^r \alpha_k^r,$$

also $j\gamma^{-1} = \alpha^{-1}(H\gamma)^{-1} t_{ii}^r$,

$$t_{ii}^r = j(H\gamma)\alpha\gamma^{-1}.$$

Es ist also t_{ii}^r der Index des betreffenden Fixpunktes, multipliziert mit einem Element der zugehörigen H -Klasse. Versteht man für $\gamma \in \Gamma$ unter $|\gamma|$ die γ enthaltende H -Klasse p_i^H und setzt für Gruppenringelemente²¹⁾ $t = \sum_i n_i \gamma_i$

$$|t| = \sum_i n_i |\gamma_i| = \sum_i n_i p_{ii}^H,$$

so stellt die Reidemeistersche Spureninvariante²²⁾

$$(15) \quad R(T) = \left| \sum_{i,r} (-1)^r t_{ii}^r \right| = \sum_i v_i p_{ii}^H$$

genau das System der H -Klassen, jede mit der zugehörigen Fixpunktzahl versehen, also den Ausdruck (9) aus § 2 dar. Wir können demnach folgendes feststellen:

Satz 2. Es sei K eine Zerlegung von \mathfrak{P} und K_1 eine Unterteilung von K ; $g(x)$ sei eine Simplicialabbildung von K_1 in K mit nur Grund-Fixsimplexen, $g(c)$ der dadurch erzeugte Automorphismus der Ketten von K_1 . Bezeichnet R_0 nach (9) das System der Fixpunktklassen bei $g(x)$, so liegt über $g(c)$ ein Automorphismus T des Homotopiekettenringes von K_1 mit

$$(16) \quad R(T) = R_0.$$

Korollar 1. Die Invariante R_0 einer Abbildung f läßt sich berechnen, indem man für eine Simplicialapproximation $g(x)$ nach Satz 2 die Spureninvariante $R(T)$ bildet; es gilt (16).

Korollar 2. Sind T_1 und T_2 Automorphismen desselben Homotopiekettenringes, die über homotopen Simplicialabbildungen der in Satz 2 bezeichneten Art liegen, so ist mit passendem $\gamma \in \Gamma$

$$R(T_2) = \gamma R(T_1).$$

In Korollar 1 wird von dem Umstand Gebrauch gemacht, daß die Simplicialapproximation $g(x)$ durch eine kleine Deformation aus $f(x)$ hervorgeht; es ist darauf zu achten, daß der richtige, d. h. dem System $\{f, \mathbb{C}\}$ entsprechende Automorphismus T genommen wird. — In Korollar 2 mußte berücksichtigt werden, daß T durch die Abbildung des Grundkomplexes nur bis auf einen

²¹⁾ Der Übergang von t zu $|t|$ im Gruppenring ist eine Restklassenbildung nach der (additiven) Untergruppe der t_0 mit $|t_0| = 0$.

²²⁾ In R. 1 mit $|s|$ bezeichnet.

inneren Automorphismus und R daher nur bis auf einen links stehenden Faktor aus Γ bestimmt ist (R. 1, S. 592).

Die Frage, ob (16) stets gilt, sobald für eine Abbildung beide Seiten Sinn haben, kann hier übergangen werden, da bereits die in Satz 2 zugelassenen Abbildungen in \mathfrak{f} überall dicht liegen. Die Identität (16) kann deshalb nicht in gleicher Allgemeinheit wie diejenige zwischen der Lefschetz-Hopfischen Spureninvarianten A und der Fixpunktgesamtzahl ausgesprochen werden, weil R zur Zeit nicht als stetigkeitstopologische Größe definiert werden kann. A dagegen ließ sich durch Homologiegruppen ausdrücken und war deshalb schon unabhängig vom Fixpunktsatz deformationsinvariant.

§ 5.

Invariante Formulierung des Ergebnisses.

Aus der in § 2, (10) gewonnenen von $\{f, \mathbb{C}\}$ abhängigen deformationsinvarianten Größe haben wir nun eine von \mathbb{C} unabhängige Homotopieinvariante abzuleiten. Es wäre unzumutbar gewesen, von vornherein nur mit Invarianten im großen zu arbeiten, da die Übersichtlichkeit der Ausdrücke und die Durchsichtigkeit der Methode hierunter gelitten hätten.

Wir wissen aus § 1: bei Ersatz von $[\mathbb{C}]$ durch $\gamma[\mathbb{C}]$ geht H in $T_\gamma H$ über²³⁾. Aus (3) entnimmt man, daß dann die den Fixpunkten zugeordneten Gruppenelemente alle links den Faktor γ erhalten. Jede H -Klasse ist also zu ersetzen durch die entsprechende $T_\gamma H$ -Klasse

$$p_i^{\gamma H} = \gamma p_i^H,$$

der dieselbe Fixpunktklasse, also auch die gleiche Fixpunktzahl zugeordnet ist. Der Ausdruck R_0 (Formel (9)) geht über in $\sum_k v_k \gamma p_k^H = \gamma R_0$.

Satz 3. Es sei \mathfrak{f} eine Klasse von Abbildungen des Polyeders \mathfrak{P} in sich, Γ die Wegegruppe von \mathfrak{P} mit dem Aufpunkt p . Wählt man eine Abbildung f aus \mathfrak{f} und einen Weg \mathbb{C} von p nach $f(p)$ und bildet nach § 2 H und R_0 , so ist

$$(17) \quad \Sigma_2 = \{T_\gamma H, \gamma R_0 \ (\gamma \in \Gamma)\}$$

eine von f und \mathbb{C} unabhängige Invariante.

Der Ausdruck (17) soll die Menge aller Paare $\{T_\gamma H, \gamma R_0\}$ bezeichnen, die erhalten wird, wenn man γ ganz Γ durchlaufen läßt. — Der Beweis wurde soeben erbracht.

Die Tatsache, daß \mathfrak{f} eine Automorphismenfamilie Σ_1 invariant zugeordnet ist, ist in Satz 3 als Teilaussage enthalten. Desgleichen lassen sich aus (17),

²³⁾ Im Falle mehrstufiger Automorphismen ist der Ausdruck $T_\gamma H$ allgemeiner als $HT_\gamma = T_{H\gamma} H$; vgl. R. 1, S. 592.

ja schon aus dem Ausdruck R_0 (Formel (9)), die Fixpunktzahlen v_k ablesen; etwa nach nicht wachsenden Absolutbeträgen geordnet und bei gleichem Betrag die positiven den negativen Zahlen vorangestellt, bilden sie eine wohlbestimmte \mathfrak{f} invariant zugeordnete Folge von Zahlen

$$(18) \quad v_1, v_2, \dots, v_r \text{ (für } r < \infty) \text{ bzw. } v_1, v_2, \dots \text{ (für } r = \infty).$$

Hier ist r die durch Γ und H oder $\{T, H (\gamma \in \Gamma)\}$ festgelegte Zahl der H -Klassen. Es ist $v_k = 0$ für $k > \mu$, falls $r > \mu$; $v_k \neq 0$ für $k \leq \mu$. Die Zahlenfolge v_1, \dots, v_μ war freilich bereits durch W. 1 als Invariante der Abbildungsklasse erkannt, doch enthält (17) natürlich mehr als dies. Durch (18) ist auch die

Gesamtzahl der Fixpunkte, die Lefschetzsche Zahl $A = \sum_{k=1}^r v_k = \sum_{k=1}^r v_k$ ausdrückbar. Aus der Zahlenfolge (18) ist ferner zu ersehen, wieviel unwesentliche Fixpunktklassen bei einer Abbildung aus \mathfrak{f} auftreten können; denn aus dem Beweis des letzten Teiles des Hilfssatzes, in dem eine leere Fixpunktklasse „sichtbar gemacht“ wird, liest man ab, daß jenes Verfahren gleichzeitig für beliebig (endlich) viele Fixpunktklassen durchführbar ist, da es nur eine Abänderung auf einem beliebig kleinen Teilbereich an beliebiger Stelle erfordert. Es folgt: Im Falle $r < \infty$ gibt es maximal $r - \mu$ unwesentliche Fixpunktklassen, im Falle $r = \infty$ enthält \mathfrak{f} Abbildungen mit beliebig vielen Fixpunktklassen. In § 8 werden hierfür Beispiele gegeben.

§ 6.

Zuordnung zwischen Fixpunktklassen verschiedener Abbildungen im großen.

Nach W. 1, Satz 2 besteht zwischen den Fixpunktklassen bei homotopen Abbildungen f und g eine eindeutige Zuordnung, die durch einen Deformationsweg vermittelt wird. Die Frage liegt nahe, ob und wann diese Zuordnung vom Wege unabhängig ist, wann also eine Fixpunktklasse bei allen Abbildungen von \mathfrak{f} individuell wiedererkennbar ist. Die Invariante \mathcal{L}_2 gibt hierüber keinen Aufschluß.

Bei solchen Fixpunktklassen, deren Fixpunktzahl in der Folge (18) nur einmal auftritt, ist es trivialerweise der Fall. Daß aber nicht stets durch den erwähnten Satz eine eindeutige Fixpunktklassenzuordnung „im großen“, unabhängig vom Deformationswege, geliefert wird, ist an folgendem einfachen Beispiel zu erkennen.

Mit ξ^* bezeichnen wir für reelles ξ die Restklasse der λ mit $\lambda \equiv \xi \pmod{1}$. \mathfrak{P} sei homöomorph zu einem Kreis; wir betrachten \mathfrak{P} als Menge der ξ^* ($0 < \xi \leq 1$). Eine Schar von Abbildungen wird gegeben durch

$$f_r(\xi^*) = \tau^* - \xi^* \quad (0 \leq \tau \leq 1);$$

es entstehen bei f , die beiden Fixpunkte

$$p_1 = \left(\frac{\tau}{2}\right)^*, \quad q_1 = \left(\frac{\tau}{2} + \frac{1}{2}\right)^*.$$

Man sieht leicht, daß p_0 und q_0 bei f_0 verschiedenen Fixpunktklassen angehören; jedoch ist $f_0 = f_1$ und der Deformationsweg $\{f_\tau \mid (0 \leq \tau \leq 1)\}$ ordnet offenbar p_0 und $p_1 = q_0$, q_0 und $q_1 = p_0$ einander zu, vertauscht also die beiden Fixpunktklassen.

Die Redeweise „in der gleichen Fixpunktklasse liegen“ ist daher, wenn es sich um verschiedene Abbildungen handelt, mit der gleichen Vorsicht zu verwenden wie der Ausdruck „dem gleichen Blatt angehören“ bei Überlagerungen. (In der Tat handelt es sich hier um eine Überlagerungsbeziehung.) Die geschlossenen bei einer festen Abbildung $f \in \mathfrak{f}$ beginnenden Deformationswege erzeugen Permutationen unter den Fixpunktklassen; diese bilden eine Gruppe g , die augenscheinlich ein homomorphes Bild der Fundamentalgruppe von \mathfrak{f} ist.

Wir fragen nun nach hinreichenden Bedingungen, unter denen die Zuordnung der Fixpunktklassen vom Wege unabhängig bleibt; oder, was auf dasselbe hinausläuft, nach notwendigen Bedingungen dafür, daß zwei bei einer Abbildung f auftretende Fixpunktklassen f_1, f_2 durch die Gruppe g transitiv verbunden²⁴⁾ sind. Diese Aussage hat einen geometrischen Sinn; es besteht dann zwischen f_1 und f_2 die in W. 1, § 3 erklärte Zuordnung.

Es sei

$$(19) \quad \{H, \sum_k v_k p_k^H\}$$

ein Wert der Invarianten (10), der durch $\{f, \mathbb{C}\}$ bestimmt ist, und g habe mindestens zwei Elemente. Durch einen bei f beginnenden geschlossenen Deformationsweg realisieren wir eine von der Identität verschiedene Permutation; die Invariante hat ihren Wert behalten, doch hat $[\mathbb{C}]$ an der Deformation teilgenommen und ist etwa in $\alpha^{-1}[\mathbb{C}]$ übergegangen, und die den p_k^H zugeordneten Fixpunktklassen sind permutiert. Ersetzen wir nun wieder $\alpha^{-1}[\mathbb{C}]$ durch $[\mathbb{C}]$, so nimmt die Invariante den Wert

$$\{T_\alpha H, \sum_k v_k \alpha p_k^H\}$$

an, der jedoch mit (19) identisch sein muß. Es folgt, daß

$$(20) \quad T_\alpha H = H$$

ist und daß unter den H -Klassen durch die Linksmultiplikation mit α eine Permutation erzeugt wird, bei der aus $\alpha p_i^H = p_i^H$ stets $v_i = v_i$ folgt. Die Bedingung (20) besagt, daß α mit jedem Element der Bildgruppe $H\Gamma$ vertauschbar ist; jedes derartige α erzeugt unter den H -Klassen eine Permutation,

²⁴⁾ Über Permutationsgruppen vgl. S., S. 110.

der aber nicht ohne weiteres eine geometrische Bedeutung zuzukommen braucht. Wir fassen das Ergebnis zusammen.

Satz 4. *Damit zwei bei einer Abbildung f auftretende Fixpunktklassen f_1, f_2 durch g transitiv verbunden sind, ist notwendig,*

- a) *daß ihre algebraischen Fixpunktzahlen gleich sind: $v_1 = v_2$;*
- b) *daß in Γ ein $\kappa \neq 1$ existiert, das mit jedem Element der Gruppe $H\Gamma$ vertauschbar ist;*
- c) *daß mit einem solchen $\kappa \in \Gamma$ $\kappa p_1^H = p_2^H$ gilt. Hierbei ist H irgendein durch f erzeugter Automorphismus von Γ .*

Die Bedingung b) ist besonders bequem anwendbar, weil sie nur die Kenntnis von Γ und H voraussetzt. Sie ist sicher nicht erfüllt, wenn folgende zwei Bedingungen gelten:

$\alpha)$ H ist einstufig, d. h. $H\Gamma = \Gamma$.

$\beta)$ Das Zentrum von Γ enthält nur das Einselement.

Bedingung $\alpha)$ ist insbesondere bei topologischen Abbildungen stets erfüllt. Beide Bedingungen sind in dem von Nielsen vorwiegend untersuchten Falle (geschlossene Flächen mit einem Geschlecht größer als 1) erfüllt. Dort ist also die Aussage, daß zwei Fixpunkte homotoper Abbildungen der gleichen Fixpunktklasse angehören, ohne weiteres sinnvoll. Im allgemeinen aber kann man höchstens die Aussage machen, daß zwei solche Fixpunktklassen dem gleichen transitiven System angehören.

§ 7.

Eine weitere Invariante.

Aus den Überlegungen von § 6 entnehmen wir, daß ein geschlossener, bei f beginnender Deformationsweg $\{f_\tau \ (0 \leq \tau \leq 1)\}$ nur dann eine von der Identität verschiedene Permutation unter den Fixpunktklassen bei f erzeugen kann, wenn das System $\{f, \mathbb{C}\}$ durch die Deformation wesentlich geändert wird. Der Bildpunkt von p beschreibe hierbei den Weg

$$(21) \quad \mathcal{W}_1 = \{f_\tau(p) \ (0 \leq \tau \leq 1)\};$$

es gehe $[\mathbb{C}]$ dadurch über in

$$(22) \quad [\mathbb{C}\mathcal{W}_1] = \kappa_1^{-1}[\mathbb{C}]$$

und die Fixpunktklassen bei f mögen die Permutation Π_1 erleiden. $[\mathcal{W}_1]$ und κ_1 bestimmen sich nach (22) gegenseitig eindeutig. Die Menge aller in (22) möglichen Elemente κ_1 heiße \mathfrak{g}_3 . Indem man in (22) einen anderen derartigen Weg \mathcal{W}_2 (mit zugehörigem κ_2, Π_2), dann \mathcal{W}_1^{-1} und $\mathcal{W}_1\mathcal{W}_2$ einsetzt, erkennt

man, daß g_3 eine Untergruppe von Γ ist, die durch die Zuordnung $\alpha \rightarrow //$ homomorph auf die Gruppe g der Permutationen $//$ abgebildet wird. Umfaßt g_4 diejenigen Elemente α_0 von g_3 , die die identische Permutation erzeugen, für die also

$$(23) \quad \alpha_0 p_k^H = p_k^H \quad (k = 1, \dots, r)$$

ist, so ist g_4 Normalteiler von g_3 , und g ist isomorph zur Faktorgruppe g_3/g_4 .

Ist g_1 die Menge der $\gamma \in \Gamma$, die die Bedingung b) in Satz 4 (oder (20)) erfüllen, so ist g_1 Obergruppe von g_3 ; bezeichnet man noch mit g_2 die Untergruppe von g_1 , deren Elemente α

$$(24) \quad \{T_\alpha H, \alpha R_0\} = \{H, R_0\}$$

erfüllen, also nur solche H -Klassen permutieren, deren Fixpunktzahlen gleich sind (Bedingung a) und c) in Satz 4), so hat man

$$(25) \quad \Gamma \supset g_1 \supset g_2 \supset g_3 \supset g_4.$$

Hier ist g_1 durch Γ und H , g_2 durch Γ , H und R_0 , g_4 und damit auch g durch Γ , H und g_3 bestimmt. Jedoch kann g_3 aus den bisher besprochenen Invarianten nicht abgelesen werden; es lassen sich Beispiele finden, aus denen ersichtlich ist, daß selbst durch Σ_2 noch nicht festgelegt wird, ob g mehr als nur ein Element enthält.

Die Gruppe g_3 , die hier durch $\{f, \mathbb{C}\}$ bestimmt wurde, ist offenbar deformationsinvariant. Um eine homotopieinvariante Aussage zu erhalten, fragen wir nach dem Einfluß einer un stetigen Abänderung von \mathbb{C} auf g_3 .

Aus (21) erkennt man, daß ein Ersatz von $[\mathbb{C}]$ durch $\gamma[\mathbb{C}]$ eine Transformation $\alpha_1 \rightarrow T_\gamma \alpha_1$ zur Folge hat. Es geht also g_3 (entsprechend übrigens auch g_1 , g_2 , g_4) in eine konjugierte Untergruppe $T_\gamma g_3$ über. Durch Kombination dieses Ergebnisses mit Σ_2 erhalten wir den invarianten Ausdruck $\{T_\gamma H, \gamma R_0, T_\gamma g_3 \mid \gamma \in \Gamma\}$.

Satz 5. Die Menge

$$(26) \quad \Sigma_3 = \{T_\gamma H, \gamma R_0, T_\gamma g_3 \mid \gamma \in \Gamma\}$$

ist eine Homotopieinvariante. g_3 ist die Menge der $\alpha_1 \in \Gamma$, für die (21) und (22) erfüllbar ist. Die Gruppe g der Permutationen der Fixpunktklassen bei f ist isomorph zu g_3/g_4 , wo g_4 die Menge der $\alpha_0 \in g_3$ mit (23) ist.

§ 8.

Beispiele.

Zur Veranschaulichung der Verhältnisse betrachten wir (wie schon in § 6) Abbildungen des Kreises in sich. \mathfrak{P} sei wieder die Menge der ξ^* mit $0 \leq \xi < 1$; $p = 0^*$; α sei die den Weg $\{\tau^* \mid (0 \leq \tau \leq 1)\}$ enthaltende Klasse. Γ ist eine

freie zyklische Gruppe mit dem erzeugenden Element α . Da Γ abelsch ist, enthält jede Automorphismenfamilie nur ein Element; es ist $g_1 = \Gamma$. Ein Automorphismus H von Γ wird völlig festgelegt durch Angabe von $H\alpha$; es gibt unendlich viele Automorphismen, sie werden definiert durch

$$H_n \alpha = \alpha^n \quad (n = 0, \pm 1, \pm 2, \dots).$$

Die Abbildung

$$f^{(n)}(\xi^*) = n \xi^*$$

erzeugt den Automorphismus H_n ; jedes $f^{(n)}$ gehört also einer besonderen Abbildungsklasse \mathfrak{t}_n an. Wir übergehen den sehr leichten Beweis, daß hiermit sämtliche Abbildungsklassen erschöpft sind. (In dem Beispiel in § 6 war $n = -1$). Für jedes n ist $f^{(n)}(p) = p$. Stets ist ein geschlossener Deformationsweg angebar, bei dem der Bildpunkt von p einen Weg der Klasse α beschreibt, nämlich $f_\tau^{(n)}(\xi^*) = f^{(n)}(\xi^*) + \tau^*$ ($0 \leq \tau \leq 1$); daher ist $g_3 = \Gamma$. Die Bedingung für die Konjugiertheit zweier Elemente α^k, α^l nach H_n lautet

$$\alpha^l = (H_n \alpha^r) \alpha^k \alpha^{-r} = \alpha^{r(n-1)} \alpha^k;$$

die H_n -Klassen sind daher die Restklassen (oder Nebengruppen) nach der von α^{n-1} oder $\alpha^{|n-1|}$ erzeugten Untergruppe $\Gamma_{|n-1|}$. Die Zahl der H -Klassen ist also

$$r_n = |n-1| \quad \text{für } n \neq 1,$$

$$r_1 = \infty.$$

Da g_3 das Element α enthält, das, als linker Faktor zu den Restklassen nach $\Gamma_{|n-1|}$ gesetzt, eine zyklische Vertauschung bewirkt, sind jeweils sämtliche Fixpunktclassen durch g transitiv verbunden. Es ist $g_4 = \Gamma_{|n-1|}$, daher ist $g = g_3/g_4$ zyklisch von der Ordnung $|n-1|$ (für $n \neq 1$) bzw. von der Ordnung ∞ (für $n = 1$).

Alle Fixpunktzahlen $v_k^{(n)}$ müssen bei festem n gleich sein, daher ist $\mu = r$ oder $\mu = 0$. Im Falle $n = 1$ (Klasse der Identität) folgt hieraus schon $\mu = 0$, da nicht $\mu = \infty$ sein kann. Es gibt also bei \mathfrak{t}_1 keine wesentlichen Fixpunktclassen; in der Tat gestattet ja der Kreis fixpunktfreie Deformationen in sich. Es gibt aber unendlich viele unwesentliche Klassen; es muß also Deformationen des Kreises mit beliebig vielen nicht leeren Fixpunktclassen geben. Solche erhält man in der Tat z. B. in der Gestalt

$$(27) \quad f_\omega^{(1)}(\xi^*) = (\xi + \omega \sin 2\pi \xi)^* \quad (\omega > 0 \text{ fest}).$$

Die (aus der Hopfschen Spurformel berechenbare) Gesamtfixpunktzahl von \mathfrak{t}^n ist $A^{(n)} = 1 - n$, daher haben wir

$$\mu = r = n - 1, \quad v_k = -1 \quad \text{für } n > 1,$$

$$\mu = r = 1 - n, \quad v_k = 1 \quad \text{für } n < 1,$$

$$\mu = 0, \quad r = \infty, \quad v_k = 0 \quad \text{für } n = 1.$$

Für $n \neq 1$ gibt es also nur $|n - 1|$ wesentliche Klassen und *keine unwesentlichen*.

Da $\tilde{\mathfrak{P}}$ die Zahlengerade ist, sind die über den f aus \mathfrak{T} liegenden Abbildungen von $\tilde{\mathfrak{P}}$ stetige reelle Funktionen

$$F(\xi) = \eta,$$

die einer Funktionalgleichung

$$F(\xi + 1) = F(\xi) + n$$

genügen. Hieraus sind die obigen Angaben über μ und ν ebenfalls leicht zu entnehmen. Insbesondere erhält man die Fixpunkte von (27) klassenweise, wenn man die über $f_w^{(1)}$ liegenden Abbildungen

$$\alpha^k F_w^{(1)}(\xi) = \xi + \omega \sin 2\pi\xi \cdot \frac{1}{k}$$

betrachtet und für verschiedene Werte von k diesen Ausdruck mit ξ gleichsetzt.

Bei diesen Beispielen war stets $\Gamma = g_1 = g_2 = g_3$. Daß der Fall $g_1 = \{1\}$ eintreten kann, d. h. daß schon g_1 nur das Einselement von Γ enthält, wurde auf S. 231 erwähnt. — Wäre stets $g_2 = g_3$, so wäre g_3 durch Σ_2 bestimmt, was, wie schon in § 7 behauptet, nicht der Fall ist. Dann und nur dann, wenn $g_2 = g_3$ ist, wenn also aus (20) (24) folgt, kann man R_0 als Funktion von f und H , insbesondere bei festem H die Fixpunktzahlen als eindeutige Funktionen der H -Klassen betrachten. Jedoch lassen sich auch für $g_2 \neq g_3$ leicht Beispiele bilden. Man kann deshalb nicht allgemein die Fixpunktzahlen den H -Klassen eindeutig zuordnen.

(Eingegangen am 16. 4. 1941.)

Über die Irreduzibilität von Polynomen.

Von

Stephan Lipka in Szeged (Ungarn).

1. Das bekannte Irreduzibilitätskriterium von O. Perron lautet folgendermaßen¹⁾:

Sind die Koeffizienten des Polynoms

$$f(x) = x^n + a_1 x^{n-1} + \dots + a_n$$

ganze rationale Zahlen und ist dabei

$$(1) \quad |a_1| > 1 + |a_2| + |a_3| + \dots + |a_n|, \\ a_n \neq 0,$$

so ist das Polynom $f(x)$ im Körper der rationalen Zahlen irreduzibel. Dieses Kriterium ist eine einfache Folgerung eines Mayerschen Satzes²⁾, nach dem das Polynom $f(x)$ $n-1$ Wurzeln innerhalb des Kreises $|x| = 1$ hat, wenn die Koeffizienten von $f(x)$ der Ungleichung (1) genügen. Der Beweis des vorigen Kriteriums folgt unmittelbar daraus, daß $f(x)$, wenn es reduzibel ist, wenigstens einen Faktor enthalten muß, dessen sämtliche Wurzeln absolut < 1 sind und der also ein konstantes Glied absolut < 1 enthält. Das konstante Glied soll aber hier, nach $a_n \neq 0$, eine von Null verschiedene ganze Zahl sein, was unmöglich ist.

L. Berwald hat folgenden Satz bewiesen³⁾: Genügen die Koeffizienten des Polynoms $f(x)$ der Bedingung

$$(1') \quad |a_1| > |1 + a_2| + |a_3| + \dots + |a_n|,$$

so hat das Polynom $f(x)$ $n-1$ Wurzeln innerhalb des Kreises $|x| = 1$. Aus diesem Satz folgt ebenso wie vorher die folgende verschärfte Form des Perronschen Kriteriums:

Genügen die Koeffizienten des Polynoms $f(x)$ der Ungleichung (1') und ist $a_n \neq 0$, so ist $f(x)$ irreduzibel.

¹⁾ O. Perron, Algebra II (1933), S. 35.

²⁾ D. E. Mayer, Sur les équations algébriques. Nouv. Ann. Math. (3) 10 (1891), S. 111–124.

³⁾ L. Berwald, Über einige mit dem Satz von Kakeya verwandte Sätze. Math. Zeitschr. 37 (1933), S. 75.

Von O. Perron stammt auch der folgende Satz⁴⁾: Wenn die Koeffizienten von $f(x)$ der Bedingung $a_2 > 0$ und außerdem der Ungleichung

$$\sqrt{a_2} > 4^{n-1} (|a_1| + |a_3| + |a_4| + \cdots + |a_n|)$$

genügen, so ist $f(x)$ irreduzibel. Ich werde zeigen, daß in der vorigen Ungleichung der Faktor 4^{n-1} sich durch einen wesentlich kleineren ersetzen läßt. Es gilt nämlich der folgende:

Satz 1. Wenn die Koeffizienten des Polynoms $f(x)$ der Bedingung $a_2 > 0$ und außerdem der Ungleichung

$$(2) \quad \sqrt{a_2} > \frac{3\sqrt{2}}{2} (|a_1| + |a_3| + |a_4| + \cdots + |a_n|), \quad a_n \neq 0,$$

genügen, so ist $f(x)$ irreduzibel.

Zum Beweis des Satzes benötigen wir den

Hilfssatz. Es sei $f(x) = x^n + a_1 x^{n-1} + \cdots + a_n$ ein Polynom mit reellen Koeffizienten, und es bedeute $-\delta < x < \delta$, $\delta > 0$, ein Intervall, das alle reellen Wurzeln von $f(x)$ enthält. Die Koeffizienten von $f(x)$ und δ sollen den folgenden Bedingungen genügen:

$$(3) \quad a_2 > 0,$$

$$(4) \quad \delta > 1,$$

$$(5) \quad a_2 > \delta^2 + (A - 1) \delta + \frac{A}{\delta},$$

wo

$$(6) \quad A = |a_1| + |a_3| + |a_4| + \cdots + |a_n|$$

ist, und

$$(7) \quad |a_1| \leq A - 1.$$

Dann hat das Polynom $f(x)$ keine reellen Wurzeln außerhalb des Intervalles $-1 < x < 1$.

Beweis. Da $\delta (> 1)$ nach Voraussetzung eine obere Grenze für die positiven Wurzeln von $f(x)$ ist, so muß das Polynom $f(x)$ für alle $x \geq 1$ von Null verschieden sein, wenn $f(x)$ für $1 \leq x \leq \delta$ nicht verschwindet. Ist nun $x > 0$, so gilt nach (3)

$$|f(x)| \geq a_2 x^{n-2} - x^n - |a_1| x^{n-1} - |a_3| x^{n-3} - \cdots - |a_n|,$$

und hieraus folgt

$$(8) \quad |f(x)| > 0,$$

wenn

$$a_2 > x^2 + |a_1| x + |a_3| \frac{1}{x} + \cdots + |a_n| \frac{1}{x^{n-2}}.$$

⁴⁾ O. Perron, Neue Kriterien für die Irreduzibilität algebraischer Gleichungen. Journ. f. d. Math. 132 (1907), S. 75.

Diese Ungleichung wird nun für $1 \leq x \leq \delta$ erfüllt, wenn

$$a_2 > x^2 + |a_1|x + \frac{|a_3| + |a_4| + \dots + |a_n|}{x} \quad (1 \leq x \leq \delta)$$

genommen wird, und nach den Bedingungen (6), (7) um so mehr erfüllt, wenn

$$a_2 > x^2 + (A - 1)x + \frac{A}{x} \quad (1 \leq x \leq \delta)$$

vorausgesetzt wird. Wächst nun x von 1 bis δ , so zeigt eine leichte Rechnung, daß die rechte Seite der vorigen Ungleichung den größten Wert für $x = \delta$ annimmt. Die Ungleichung (8) ist also für alle $x \geq 1$ sicher erfüllt, wenn a_2 der Bedingung (5) genügt. Wir müssen noch zeigen, daß das Polynom $f(x)$ auch für alle $x \leq -1$ von Null verschieden ist. Setzt man hierzu in dem Polynom $-x$ statt x (> 0) und betrachtet das normierte Polynom $(-1)^n f(-x)$ statt $f(x)$, so kann man, da der Koeffizient von x^{n-2} wieder positiv ist, so wie in den vorigen Überlegungen schließen, daß für alle $x \leq -1$ die Ungleichung (8) gilt. Damit ist der Hilfssatz bewiesen.

Beweis des Satzes 1. Für den Fall $n = 2$ ist der Satz richtig; es sei also

$$(9) \quad n > 2.$$

Nach (9) und der Ungleichung $|a_n| \geq 1$ folgt

$$(10) \quad A = |a_1| + |a_3| + |a_4| + \dots + |a_n| \geq 1$$

und hieraus nach (2)

$$(11) \quad a_2 > 1 + A,$$

da $a_2 > \frac{9}{2} A^2 \geq \frac{9}{2} A > 1 + A$ gilt. Das Polynom hat also zufolge (11) (nach Meyers Satz) $n - 2$ Wurzeln innerhalb des Einheitskreises. Zwei Wurzeln von $f(x)$ liegen also außerhalb des Kreises $|x| = 1$, und wir behaupten, daß die beiden Wurzeln konjugiert-komplexe Zahlen sind. Bedeutet nämlich H den absoluten Betrag des absolut größten negativen Koeffizienten von $f(x)$, so ist $1 + H$ eine obere Grenze der positiven Wurzeln des Polynoms $f(x)^5$, und ebenso ist $1 + K$ eine obere Grenze für absolute Beträge der negativen Wurzeln, wenn K den absoluten Betrag der kleinsten negativen Koeffizienten des normierten Polynoms $(-1)^n f(-x)$, ($x > 0$), bedeutet. Da in den Polynomen $f(x)$ und $(-1)^n f(-x)$ die Koeffizienten a_2 bzw. $(-1)^{2n-2} a_2$ positiv sind, so gilt offenbar:

$$\text{Max}\{H, K\} \leq |a_1| + |a_3| + |a_4| + \dots + |a_n| = A.$$

⁵⁾ Vgl. z. B. Bieberbach-Bauer, Algebra (1928), S. 129.

Die reellen Wurzeln von $f(x)$ liegen also sämtlich im Intervall $-(1+A) < x < 1+A$. Man wende nun den vorigen Hilfssatz für $\delta = 1+A$ an. Die Bedingungen (3), (4) sind hier selbstverständlich erfüllt. (5) ist auch erfüllt, da nach (2) $a_2 > \frac{9}{2}A^2$ ist und außerdem wegen (7) die rechte Seite von (5) (für $\delta = 1+A$) der Ungleichung

$$(12) \quad \frac{9}{2}A^2 \geq (1+A)^2 + (A^2 - 1) + \frac{A}{1+A}$$

genügt. Die Richtigkeit von (7) folgt nach (9) und der Voraussetzung $|a_n| \geq 1$. Das Polynom $f(x)$ hat also keine reellen Wurzeln außerhalb des Kreises $|x| = 1$, und im Äußeren des Kreises $|x| = 1$ liegt ein Paar konjugiert-komplexer Wurzeln.

Wenn wir nun annehmen, daß das Polynom $f(x)$ reduzibel ist und etwa in zwei Faktoren zerfällt, so muß einer der Faktoren — es sei der erste — die beiden, außerhalb des Kreises $|x| = 1$ liegenden konjugiert-komplexen Wurzeln enthalten. Der zweite Faktor soll dann lauter Wurzeln haben, deren absolute Beträge kleiner als 1 sind. Dies ist aber nach obigem unmöglich. Die Annahme der Reduzibilität war also falsch.

Wir bemerken nachträglich, daß beim Beweis des Satzes 1 die rechte Seite von (12) auch der Ungleichung

$$2(1+A)^2 \geq (1+A)^2 + (A^2 - 1) + \frac{A}{1+A}$$

genügt und deshalb die Voraussetzung (2) sich durch folgende ersetzen läßt:

$$\sqrt{a_2} > \sqrt{2}(1 + |a_1| + |a_3| + \cdots + |a_n|).$$

Die letztere Voraussetzung ist weniger weitgehend als (2), wenn $A > 2$ ist; $A > 2$ besteht z. B. dann, wenn das Polynom $f(x)$ mindestens fünf von Null verschiedene Koeffizienten hat.

Satz 2: *Das Polynom $f(x)$ habe nun eine Lücke, und zwar sei*

$$a_1 = 0.$$

Genügen die Koeffizienten von $f(x)$ den Ungleichungen

$$a_2 > 0, \quad a_n \neq 0,$$

$$(13) \quad \sqrt{a_2} > \sqrt{3}(|a_3| + |a_4| + \cdots + |a_n|)^{\frac{1}{2}},$$

so ist $f(x)$ irreduzibel.

Beweis. Wir werden den folgenden Laguerreschen Satz anwenden⁶⁾: Ist in der Gleichung $a_0x^n + a_1x^{n-1} + \cdots + a_n = 0$ mit reellen Koeffizienten a_i ,

⁶⁾ Vgl. z. B. E. Pascal, *Repertorium der höheren Analysis*, Bd. I (Aufl. 2), S. 352. Leipzig u. Berlin 1910.

der erste der negativen Koeffizienten, also a_0, a_1, \dots, a_{j-1} , positiv und bedeutet H den absoluten Betrag des absolut größten negativen Koeffizienten, so ist

$$1 + \frac{H}{a_0 + a_1 + \dots + a_{j-1}}$$

eine obere Grenze für die positiven Gleichungswurzeln. Aus diesem Satz ergibt sich, da $a_1 = 0$ und $a_2 > 0$ ist, die obere Grenze $1 + \frac{H}{1+a_2}$ für die positiven Wurzeln von $f(x)$. Bedeutet ferner K den absoluten Betrag der kleinsten negativen Koeffizienten des Polynoms $(-1)^n \cdot f(-x)$, ($x > 0$), so ist $1 + \frac{K}{1+a_2}$ eine obere Grenze für die absoluten Beträge der negativen Wurzeln von $f(x)$. Setzt man

$$A = |a_3| + |a_4| + \dots + |a_n|,$$

so gilt offenbar

$$\text{Max}\{H, K\} \leq A,$$

und infolgedessen liefert der Wert $1 + \frac{A}{1+a_2}$ eine obere Grenze für die reellen Wurzeln von $f(x)$. Wendet man nun den Hilfssatz für $\delta = 1 + \frac{A}{a_2}$ an, so gewinnt man für (5)

$$a_2 > \left(1 + \frac{A}{a_2}\right)^2 + (A-1)\left(1 + \frac{A}{a_2}\right) + \frac{A}{1 + \frac{A}{a_2}}.$$

Dies ist aber sicher erfüllt, da nach (13) $\frac{A}{a_2} < \frac{1}{3}$ besteht und man außerdem $A \geq 1$ annehmen kann. Das Polynom $f(x)$ hat also keine reellen Wurzeln außerhalb und (nach Meyers Satz) $n-2$ Wurzeln innerhalb des Kreises $|x| = 1$, weil $a_2 > 3A > A+1$ ist. Ebenso wie beim Beweis des Satzes 1 schließt man nun, daß das Polynom $f(x)$ irreduzibel ist.

Satz 3. Für den Grad n von $f(x)$ gelte nun die Voraussetzung

$$n > 4.$$

Genügen noch die Koeffizienten von $f(x)$ der Bedingung $a_4 > 0$ und außerdem der Ungleichung

$$(14) \quad \sqrt[4]{a_4} > \sqrt[4]{2} (1 + |a_1| + |a_2| + |a_3| + |a_5| + \dots + |a_n|), \quad a_n \neq 0,$$

so kann $f(x)$ höchstens in zwei Faktoren zerfallen.

Beweis. Wir zeigen, daß das Polynom $f(x)$ keine reellen Wurzeln außerhalb des Kreises $|x| = 1$ hat. Man setze

$$A = |a_1| + |a_2| + |a_3| + |a_5| + \dots + |a_n|.$$

Da in den Polynomen $f(x)$ und $(-1)^n \cdot f(-x)$ die Koeffizienten a_4 bzw. $(-1)^{2n-4} a_4$ positiv sind, so sollen die sämtlichen reellen Wurzeln von $f(x)$ im Intervall $-(1+A) < x < 1+A$ liegen. Das Polynom $f(x)$ hat also keine reellen Wurzeln außerhalb des Kreises $|x| = 1$, wenn $|f(x)| > 0$ für $1 \leq x \leq 1+A$ ist. Die letztere Forderung ist aber erfüllt, wenn

$$a_4 > x^4 + |a_1|x^3 + |a_2|x^2 + |a_3|x + |a_5|\frac{1}{x} + \dots + |a_n|\frac{1}{x^{n-4}} \\ (1 \leq x \leq 1+A)$$

gilt, und nach den Beziehungen $|a_1| \leq A$, $|a_2| \leq A$, $|a_3| \leq A$ um so mehr, wenn

$$(15) \quad a_4 > (1+A)^4 + A\{(1+A)^3 + (1+A)^2 + (1+A)\} + A$$

gesetzt wird. Es gilt aber die Ungleichung

$$2(1+A)^4 > (1+A)^4 + A\{(1+A)^3 + (1+A)^2 + (1+A) + 1\},$$

und infolgedessen, weil nach (14) $a_4 > 2(1+A)^4$ ist, ist die Forderung (15) erfüllt. Das Polynom $f(x)$ hat also keine reellen Wurzeln außerhalb des Kreises $|x| = 1$. Nach (14) folgt für a_4 die Ungleichung $a_4 > 1+A$, das Polynom $f(x)$ hat also $n-4$ Wurzeln innerhalb des Kreises $|x| = 1$ (nach Meyers Satz). Im Äußeren des Kreises $|x| = 1$ liegen also zwei konjugiert-komplexe Wurzelpaare von $f(x)$. Wenn wir nun annehmen, daß das Polynom $f(x)$ in drei Faktoren zerfällt, so muß einer der Faktoren lauter Wurzeln haben, deren absolute Beträge < 1 sind. Dies ist aber unmöglich. Das Polynom kann also höchstens in zwei Faktoren zerfallen.

2. Aus dem Meyerschen Satz lassen sich noch die folgenden Irreduzibilitätskriterien ableiten:

Satz 4. Sind a_0, a_1, \dots, a_n feste ganze Zahlen und $a_1 \neq 0$, so ist

$$f(x) = a_0 p^k + a_1 x + \dots + a_n x^n$$

irreduzibel, falls p eine Primzahl und k eine genügend große ganze positive Zahl bedeutet.

Beweis. Wenn wir annehmen, daß das Polynom $f(x)$ reduzibel ist und etwa in zwei Faktoren zerfällt, so zerfällt das zu $f(x)$ reziproke Polynom auch in zwei Faktoren. Es ist also

$$(16) \quad a_0 p^k x^n + a_1 x^{n-1} + \dots + a_n \\ = (\alpha_0 p^x x^r + \alpha_1 x^{r-1} + \dots + \alpha_r) (\beta_0 p^{k-x} x^s + \beta_1 x^{s-1} + \dots + \beta_t).$$

Dabei kann $x \leq k-x$ angenommen werden. Bedeutet p^m die größte Potenz von p , durch welche der Koeffizient $a_1 \neq 0$ teilbar ist, so folgt nach der Formel

$$a_1 = \alpha_0 p^x \beta_1 + \alpha_1 \beta_0 p^{k-x}$$

die Ungleichung $\kappa \leq m$. Hieraus folgt, daß das Wurzelprodukt $\frac{\alpha_r}{\alpha_0 p^{\kappa}}$ des ersten Faktors von (16) absolut immer größer als $\frac{1}{|\alpha_0| p^m}$ ist. Nach Mayers Satz werden aber die absoluten Beträge der Wurzeln von (16) beliebig klein, wenn nur k eine genügend große ganze positive Zahl bedeutet. Die Zerlegung (16) ist also unmöglich.

Satz 5. Sind a_0, a_1, \dots, a_n feste ganze Zahlen und ist p eine genügend große Primzahl, so ist

$$f(x) = a_0 p + a_1 x + a_2 x^2 + \dots + a_n x^n$$

irreduzibel.

Beweis. Ist p genügend groß, so liegen alle Wurzeln von $f(x)$ außerhalb des Kreises $|x| = |\alpha_0|$. Wenn nun das Polynom $f(x)$ in zwei Faktoren zerfällt, so muß einer der Faktoren ein konstantes Glied haben, dessen Betrag $\leq |\alpha_0|$ ist, was unmöglich ist.

Dieser Satz ist sonst eine unmittelbare Folgerung eines Satzes von O. Ore⁷⁾.

Aus den Sätzen 4 und 5 folgt leicht der folgende Satz von L. Weisner⁸⁾: $f(x)$ sei ein ganzzahliges Polynom n -ten Grades und c bezeichne eine feste ganze Zahl. Hat nun das Polynom $f(x)$ mindestens eine rationale Wurzel $\frac{\alpha}{\beta}$ und ist p eine genügend große Primzahl, so ist das Polynom $f(x) + cp$ irreduzibel. Ist ferner $f'(\frac{\alpha}{\beta}) \neq 0$, p eine Primzahl und k eine genügend große ganze Zahl, so ist das Polynom $f(x) + cp^k$ irreduzibel.

Man setze zum Beweis $x = \frac{\alpha}{\beta} + y$ und bilde die Gleichung

$$f(x) + cp^k = f\left(\frac{\alpha}{\beta} + y\right) + cp^k.$$

Multipliziert man die beiden Seiten dieser Gleichung mit β^n , so gewinnt man das ganzzahlige Polynom

$$\beta^n (f(x) + cp^k) = b_0 y^n + b_1 y^{n-1} + \dots + b_{n-1} y + \beta^n c p^k.$$

Man kann nun für dieses Polynom den Satz 4 anwenden.

3. Definition. Es sei $f(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_n$ ein Polynom mit reellen Koeffizienten. In einer Ebene mit einem Descartesschen Koordinatensystem betrachte ich das Polygon mit den Eckpunkten:

$$(-1, 0), (0, a_0), (1, a_1), (2, a_2), \dots, (n, a_n), (n+1, 0).$$

Dieses Polygon nenne ich das Koeffizientenpolygon des Polynoms $f(x)$.

⁷⁾ O. Ore, Einige Bemerkungen über Irreduzibilität. Jahresbericht d. DMV, 44 (1934), S. 151, Satz 5.

⁸⁾ L. Weisner, Criteria for the irreducibility of polynomials. Bulletin of the American Math. Soc. 40 (1934), S. 864–870.

Satz 6. Sind die Koeffizienten des Polynoms $f(x) = x^n + a_1 x^{n-1} + \dots + a_n$ ganze positive Zahlen und ist das Koeffizientenpolygon von $f(x)$ im Punkte $(1, a_1)$ konkav (von unten), in den anderen Eckpunkten konvex, so ist $f(x)$ irreduzibel.

Beweis. Vom Verfasser rührt der folgende Satz her⁹⁾:

Es sei $g(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_n$ ein Polynom mit reellen positiven Koeffizienten. Wenn das Koeffizientenpolygon von $g(x)$ in dem Eckpunkt (k, a_k) konkav (von unten) und in den anderen Eckpunkten konvex ist, so hat das Polynom $g(x)$ $n - k$ Wurzeln innerhalb und k Wurzeln außerhalb des Kreises $|x| = 1$.

Wendet man diesen Satz auf $f(x)$ an, so ergibt sich, daß $n - 1$ Wurzeln von $f(x)$ innerhalb des Kreises $|x| = 1$ liegen. Nach dem Perronschen Prinzip ist also das Polynom irreduzibel.

Beispiel. Es sei $f(x) = x^4 + 10x^3 + 6x^2 + 3x + 1$. Das Polynom $f(x)$ ist nach Satz 6 irreduzibel. Die Irreduzibilität von $f(x)$ folgt hier aus dem Perronschen Kriterium nicht, da $|a_1| < 1 + |a_2| + |a_3| + |a_4|$ ist.

4. E. L. Petterson benutzte mit Erfolg die folgende Irreduzibilitätsbedingung¹⁰⁾: Liegen $n - 1$ Wurzeln des ganzzahligen Polynoms $f(x)$ des Grades n außerhalb des Kreises $|x| = \varrho$ und ist $0 < f(0) \leq \varrho$, so ist $f(x)$ irreduzibel. Durch Anwendung dieser Bedingung hat Petterson unter anderem den folgenden Satz bewiesen:

Es sei $f(x)$ ein normiertes Polynom der Form¹¹⁾

$$f(x) = g(x)x + M(x),$$

wo $g(x)$ und $M(x)$ ganzzahlige Polynome sind. $f(x)$ ist dann irreduzibel, falls es eine Zahl ϱ gibt, die folgende Bedingungen erfüllt:

1. $g(x)$ hat keine Wurzeln innerhalb des Kreises $|x| = \varrho$.
2. $|g(x)| > \frac{M(x)}{\varrho}$ längs des Kreises $|x| = \varrho$.
3. $1 \leq |M(0)| \leq \varrho$.

Wir beweisen nun die folgende Ergänzung dieses Satzes:

Satz 7. Es sei $f(x)$ ein Polynom der Form

$$f(x) = g(x)x + M(x),$$

⁹⁾ St. Lipka, Zur Theorie der algebraischen Gleichungen mit positiven Koeffizienten. Acta litt. ac Sc. Math. 5 (1931), S. 76, Satz V.

¹⁰⁾ E. L. Petterson, Einige aus den Größenbeziehungen der Wurzeln abgeleitete Irreduzibilitätskriterien. Math. Annalen 114 (1937), S. 79.

¹¹⁾ Wir bemerken, daß die Voraussetzung, laut der $f(x)$ ein normiertes Polynom ist, eigentlich überflüssig ist.

wo $g(x)$ und $M(x)$ ganzzahlige Polynome sind. $f(x)$ ist dann irreduzibel, falls es eine Zahl ϱ gibt, die folgenden Bedingungen erfüllt:

1. $g(x)$ hat keine Wurzeln innerhalb des Kreises $|x| = \varrho$, bis auf die Stelle $x = 0$, wo $g(0) = 0$ ist, aber $g'(0) \neq 0$.
2. $|g(x)| > \frac{|M(x)|}{\varrho}$ längs des Kreises $|x| = \varrho$.
3. $M(x)$ hat keine reellen Wurzeln innerhalb des Intervalles $-\varrho \leq x \leq \varrho$.
4. $\operatorname{sgn} M(0) = \operatorname{sgn} g'(0)$.
5. $1 \leq |M(0)| \leq \varrho$.

Beweis. Nach der Voraussetzung 1. liegen zwei Wurzeln des Polynoms $xg(x)$ innerhalb des Kreises $|x| = \varrho$. Nach Voraussetzung 2. besitzt also $f(x)$ gleichfalls zwei Wurzeln innerhalb des Kreises $|x| = \varrho$ (laut des Rouchéschen Satzes). Wegen 1., 3. und 4. hat aber das Polynom $f(x)$ keine reellen Wurzeln innerhalb des Intervalles $-\varrho \leq x \leq \varrho$. Infolgedessen hat $f(x)$ ein Paar konjugiert-komplexer Wurzeln innerhalb des Kreises $|x| = \varrho$. Wenn wir nun annehmen, daß das Polynom $f(x)$ in zwei Faktoren zerfällt, so muß einer der Faktoren lauter Wurzeln haben, deren absolute Beträge größer als ϱ sind. Das konstante Glied dieses Faktors ist dann absolut größer als ϱ , was der Annahme 5. offenbar widerspricht.

Als Beispiel erwähne ich das ganzzahlige Polynom

$$f(x) = (x^m + a)x^2 + x^m + b,$$

wo $g(x) = (x^m + a)x$ und $M(x) = x^m + b$ ist. Dies ist irreduzibel, wenn m gerade, $a \geq 0$, $b > 0$ und

$$\frac{a+b}{2} > b^m$$

ist.

Aus Satz 7. und Überlegungen von Petterson [vgl. Fußnote 10)] lassen sich noch die folgenden Sätze ableiten:

Satz 8. Es sei $f(x)$ ein ganzzahliges Polynom der Form

$$f(x) = g(x)x + a, \quad a \neq 0.$$

Die sämtlichen Wurzeln von $g(x)$ seien reell, ferner sei $x = 0$ eine einfache Wurzel von $g(x)$ und

$$\operatorname{sgn} g'(0) = \operatorname{sgn} a.$$

Ist dann ϱ der absolute Betrag der absolut kleinsten reellen Wurzel der Gleichung

$$|g(x)| = |x|,$$

so ist $f(x)$ irreduzibel für

$$|a| < \varrho.$$

Der Satz 8 läßt sich auf das Polynom $f(x) = x^2 \prod_{v=1}^n (x - b_v) + a$ mit ganzen b_v und a anwenden:

Satz 9. *Das Polynom*

$$f(x) = x^2 \prod_{v=1}^n (x - b_v) + a, \quad a \neq 0,$$

wo a und alle b_v ganz sind und

$$1 < |b_1| \leq |b_2| \leq \dots \leq |b_n|,$$

ferner

$$\operatorname{sgn} \prod_{v=1}^n b_v = \operatorname{sgn} a$$

ist, ist irreduzibel für

$$|a| < |b_1| - 1.$$

5. Mit Hilfe der Pettersonschen Irreduzibilitätsbedingung und meines Satzes [vgl. Fußnote 9)] lassen sich noch die folgenden Irreduzibilitätssätze ableiten:

Satz 10. *Genügen die Koeffizienten des ganzzahligen Polynoms $f(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_n$ der Ungleichung*

$$|a_{n-1}| > |a_0| |a_n|^{n-1} + |a_1| |a_n|^{n-2} + |a_2| |a_n|^{n-3} + \dots + |a_{n-2}| |a_n| + 1, \quad a_n \neq 0,$$

so ist $f(x)$ irreduzibel. Der Beweis folgt aus der Irreduzibilitätsbedingung von Pettersen für $\varrho = |a_n|$ mit Hilfe des Mayerschen Satzes.

Satz 11. *Genügen die Koeffizienten von $f(x)$ den Ungleichungen*

$$a_{n-2} > 0, \quad a_n > 0,$$

$$a_{n-2} a_n > A^2 \{ |a_0| a_n^{n-1} + |a_1| a_n^{n-2} + |a_2| a_n^{n-3} + \dots + |a_{n-3}| a_n^2 + |a_{n-1}| + 1 \},$$

wo

$$A = |a_0| + |a_1| + \dots + |a_{n-3}| + |a_{n-1}| + |a_n|,$$

ist, so ist $f(x)$ irreduzibel. Man beweist genau so wie beim Satz 1, daß das Polynom $f(x)$ $n-2$ Wurzeln außerhalb und zwei konjugiert-komplexe Wurzeln innerhalb des Kreises $|x| = a_n$ hat. Nach der vorigen Irreduzibilitätsbedingung folgt also, daß $f(x)$ irreduzibel ist.

Satz 12. *Sind die Koeffizienten des Polynoms $f(x) = a_0x^n + a_1x^{n-1} + \dots + a_n$ ganze positive Zahlen und genügen sie den Ungleichungen*

$$a_n < \frac{a_1}{2a_0}, \quad a_n < \frac{a_2}{2a_1}, \quad \dots, \quad a_n < \frac{a_{n-2}}{2a_{n-3}}, \quad a_n < \frac{a_{n-1}}{2a_{n-2}},$$

so ist $f(x)$ irreduzibel.

Beweis. Wir wenden die Pettersonsche Irreduzibilitätsbedingung für $\varrho = a_n$ an und beweisen, daß das Koeffizientenpolygon des Polynoms $f(a_n x)$ im Eckpunkte $(n-1, a_{n-1}a_0)$ konkav und in anderen Eckpunkten konvex ist. Es gelten wirklich die Ungleichungen

$$a_k a_n^{n-k} < \frac{a_{k-1} a_n^{n-k+1} + a_{k+1} a_n^{n-k-1}}{2}$$

für $k = 0, 1, 2, \dots, n-2$, wo $a_\lambda = 0$ für $\lambda < 0$ ist, da nach Voraussetzung $(a_{k+1} - 2a_k a_n) a_n^{n-k-1}$ positiv sind. Laut meinem Satze hat also das Polynom $f(x)$ eine Wurzel innerhalb und $n-1$ Wurzeln außerhalb des Kreises $|x| = a_n$. Da $f(0) = a_n$ ist, folgt nach der vorigen Irreduzibilitätsbedingung, daß $f(x)$ irreduzibel ist.

(Eingegangen am 23. 12. 1940.)

Eine neue Einteilung der Permutationen.

Von

Lothar v. Schrutka in Wien.

1. *Aufstiege und Abstiege. Profil einer Permutation.* Ist $ab \dots ik \dots m$ eine Permutation der Zahlen $1, 2, \dots, n$, so möge gesagt werden, daß zwei nebeneinander stehende Zahlen (Elemente) i und k einen *Aufstieg* oder einen *Abstieg* bilden, je nachdem $i < k$ oder $i > k$ ist. Die Anzahl der Aufstiege und der Abstiege zusammen beträgt $n - 1$. Ein Aufstieg möge durch das Zeichen $/$, ein Abstieg durch das Zeichen \backslash angedeutet werden. Jeder Permutation entspricht daher eine Aufeinanderfolge von $n - 1$ Zeichen $/$ oder \backslash , die etwa das *Profil* der Permutation genannt werden möge. Dagegen entsprechen einem solchen Profil im allgemeinen mehrere Permutationen, wie schon daraus ersichtlich ist, daß es $n!$ Permutationen, dagegen 2^{n-1} Profile gibt, und daß zwar für $n = 1$ und 2 $n! = 2^{n-1}$, dagegen für $n \geq 3$ $n! > 2^{n-1}$ ist. So z. B. entspricht der Permutation 1675243 das Profil $// \backslash \backslash /$, umgekehrt diesem Profil nicht nur die Permutation 1675243, sondern auch 1453276, 1276354 und noch andere.

Permutationen, die Auf- und Abstiege in regelmäßigem Wechsel aufweisen, sind von D. André in Journ. de Math. (3) 7 (1881), S. 167 und Paris C. R. 97 (1883), S. 983, 1356 betrachtet und *alternierende* genannt worden; er hat einen bemerkenswerten Zusammenhang mit den Sekanten- und den Tangentenkoeffizienten aufgedeckt (auch dargestellt in E. Netto, Lehrbuch der Kombinatorik, § 60, 63). Solche alternierenden Permutationen sind z. B. 25143, 51324 usw.

2. *Aufstiegzüge und Abstiegzüge.* Eine Gruppe von Elementen mit lauter Aufstiegen, die beiderseits von einem Abstieg eingeschlossen werden oder bis an den Rand reichen, bildet einen *Aufstiegzug*. Hierunter soll auch der Fall von 0 Aufstiegen eingeschlossen werden. Ganz entsprechend ist ein *Abstiegzug* zu erklären. Die Anzahl der Aufstiegzüge und der Abstiegzüge zusammen beträgt $n + 1$. In der Permutation 1675243 bilden also 167, 5, 24, 3 Aufstiegzüge, 1, 6, 752, 43 Abstiegzüge.

Das letzte Element eines Aufstiegzuges bildet mit dem ersten des folgenden einen Abstieg, und umgekehrt bilden die beiden Elemente eines Abstiegs das Ende und den Anfang je eines Aufstiegzuges. Daher ist die Anzahl der Aufstiegzüge um 1 größer als die der Abstiege. Ebenso ist die Anzahl der Abstiegzüge um 1 größer als die der Aufstiege. Bei Anzahlbestimmungen

kann man sich daher auf die Auf- oder Abstiege beschränken. Im folgenden sollen in erster Linie die Abstiege gezählt werden.

Die Auf- und Abstiegszüge sind von J. Bienaymé [Paris C. R. 81 (1875), S. 418] unter dem Namen *steigende* und *fallende Sequenzen* behandelt worden (auch dargestellt in E. Netto, Lehrbuch der Kombinatorik, § 60).

3. *Kennzeichnung der Profile durch Zahlen.* Setzt man in einem Profil für / die Ziffer 0 und für \ die Ziffer 1 ein und liest die Zahl im Zweiersystem, so bekommen alle möglichen Profile fortlaufende Zahlen von 0 bis $2^{n-1} - 1$.

4. *Einteilung der Permutationen. Die Anzahlen C_k^n .* Man kann die Permutationen nach der Zahl der Abstiege einteilen. Es sei C_k^n die Anzahl der Permutationen von n Elementen mit k Abstiegen (daher mit $n - 1 - k$ Aufstiegen). Schreibt man alle Permutationen in umgekehrter Anordnung, so wird aus jedem Aufstieg ein Abstieg und umgekehrt. Also ist $C_k^n = C_{n-1-k}^n$. Hierbei muß $n \geq 2$ sein.

Als Beispiel seien die 24 Permutationen für $n = 4$ angeführt:

Permutation	Profil	Zahl		Anzahl der Abstiege	Permutation	Profil	Zahl		Anzahl der Abstiege
		dya- disch	deka- disch				dya- disch	deka- disch	
1234	/	000	0	0	3124	/	100	4	1
1243	/	001	1	1	3142	/	101	5	2
1324	/	010	2	1	3214	/	110	6	2
1342	/	001	1	1	3241	/	101	5	2
1423	/	010	2	1	3412	/	010	2	1
1432	/	011	3	2	3421	/	011	3	2
2134	/	100	4	1	4123	/	100	4	1
2143	/	101	5	2	4132	/	101	5	2
2314	/	010	2	1	4213	/	110	6	2
2341	/	001	1	1	4231	/	101	5	2
2413	/	010	2	1	4312	/	110	6	2
2431	/	011	3	2	4321	/	111	7	3

Zählt man ab, so findet man für die Permutationen mit 0, 1, 2, 3 Abstiegen die Anzahlen 1, 11, 11, 1. Bei $n = 3$ findet man in ähnlicher Weise für die Permutationen mit 0, 1, 2 Abstiegen die Anzahlen 1, 4, 1. Bei $n = 2$ gibt es nur die beiden Permutationen 12 und 21; die erste hat keinen, die zweite einen Abstieg. Bei $n = 1$ gibt es nur die eine Permutation 1, die keinen Abstieg hat.

5. *Rücklaufformel für die Zahlen C_k^n .* Für höhere n wäre die Auszählung schon sehr mühsam. Die C_k^n lassen sich aber rücklaufend (rekurrent) berechnen. Zur Herleitung dient ein Verfahren, das auch von D. André angewendet wird. Wird in einer Permutation von 1, 2, ..., n die Zahl n gestrichen, so bleibt eine Permutation der Zahlen 1, 2, ..., $n - 1$ übrig. Stand n am linken Rand, so fällt im Profil das erste Zeichen \ weg; stand n im Innern, so wird ein

Zeichenpaar \diagdown durch \diagup oder durch \diagdown ersetzt, je nach der Größenbeziehung der beiden Elemente, die dem n benachbart waren; stand endlich n am rechten Rand, so fällt das letzte Zeichen \diagup weg. Im ersten Fall wird die Anzahl der Abstiege um eins kleiner; im zweiten Fall wird sie um eins kleiner oder sie bleibt unverändert, je nachdem \diagup oder \diagdown an Stelle von \diagup tritt; im dritten Fall bleibt sie unverändert. Also werden aus jeder der C_k^n Permutationen von $1, 2, \dots, n$ mit k Abstiegen so viele Permutationen von $1, 2, \dots, n-1$ mit $k-1$ Abstiegen, als sie selbst Aufstiege haben und noch eine dazu, somit $n-1-k+1=n-k$, und so viele Permutationen von $1, 2, \dots, n-1$ mit k Abstiegen, als sie selbst Abstiege haben und noch eine dazu, somit $k+1$.

Also gilt die Rücklauf- (Rekursions-) formel

$$C_k^n = n - k \overline{C_{k-1}^{n-1}} + k + 1 \overline{C_k^{n-1}}.$$

Nach dieser lassen sich die Zahlen C leicht berechnen.

6. Werte der Zahlen C_k^n . Die ersten Werte enthält die folgende Tabelle:

$n \backslash k$	0	1	2	3	4	5
1	1					
2	1	1				
3	1	4	1			
4	1	11	11	1		
5	1	26	66	26	1	
6	1	57	302	302	57	1
7	1	120	1191	2416	1191	120
8	1	247	4293	15619	15619	4293
9	1	502	14608	88234	156190	88234
10	1	1013	47840	455192	1310354	1310354

In gewissem Sinne ist eine Anordnung im Dreieck, wie bei den Binomialkoeffizienten, anschaulicher:

			1			
			1		1	
		1		4		1
	1		11		11	
1		26		66		26
						1

.....

Nach der Bedeutung der Zahlen C_k^n ist die Summe der Zahlen in jeder Zeile gleich $n!$. Dasselbe folgt durch vervollständigte Induktion aus der Rücklaufformel.

7. Auftreten der Zahlen C_k^n als höhere Differenzen. Die Zahlen C_k^n treten an verschiedenen Stellen in der Mathematik auf. Sie finden sich zuerst bei

P. S. Laplace, Mém. de l'acad. des sciences à Paris 1777, S. 99 usw., und bei S. F. Lacroix, Traité des différences, t. 3, 2^{de} éd. 1800, S. 114 (beides angeführt nach L. Saalschütz, Vorlesungen über die Bernoullischen Zahlen, Berlin, Springer, 1893, S. 65, Anm. 1, mit einigen Ergänzungen nach J. C. Poggendorff, Lit.-biogr. Handwörterbuch, Band 1; bei Saalschütz ist $C_k^n = A_{k+1}^n$). Die Formel von Laplace lautet [Saalschütz, Vorl., S. 63, Formel (32); S. 76, Formel (12)]:

$$C_k^n = (k+1)^n - \binom{n+1}{1} k^n + \binom{n+1}{2} (k-1)^n - \dots + (-1)^n \binom{n+1}{k} 1^n.$$

Sie kann sogleich durch Nachweis der Rücklaufformel in 5. bestätigt werden (L. Saalschütz, Vorl., S. 63, 64).

Für die Laplacesche Formel kann noch eine recht anschauliche Darstellung gegeben werden. Sie unterscheidet sich nämlich von der Formel für die $(n+1)$ -te Differenz für die Zahlenfolge der n -ten Potenzen der ganzen Zahlen

$$(a+n+1)^n - \binom{n+1}{1} (a+n)^n + \binom{n+1}{2} (a+n-1)^n - \dots + (-1)^{n+1} a^n,$$

wo $a = -n, -n+1, \dots, -1$ zu setzen ist, nur dadurch, daß an Stelle der n -ten Potenzen der negativen Zahlen 0 eintritt. Die Zahlen C_k^n sind daher die $(n+1)$ -ten Differenzen der Zahlenfolge $\dots, 0, 0, 0, 0, 1^n, 2^n, 3^n, 4^n, \dots$, soweit sie nicht Null sind. Der Fall $n=4$ sei hier mit Hilfe des Differenzenspiegels dargestellt:

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	1
0	0	0	0	1	11
1	1	13	12	11	
16	15	50	36	23	1
81	65	110	60	24	0
256	175	194	84	24	0
625	369	302	108	24	0
1296	671	434	132	24	
2401	1105				

8. Auftreten der Zahlen C_k^n als Entwicklungskoeffizienten in höheren Ableitungen. Die Zahlen C_k^n treten ferner, mit abwechselnden Vorzeichen, als Koeffizienten der Zähler der Ausdrücke

$$\frac{v}{(1+v)^2}, \left(\frac{v}{(1+v)^2}\right)', \left(v \left(\frac{v}{(1+v)^2}\right)'\right)', \left(v \left(v \left(\frac{v}{(1+v)^2}\right)'\right)'\right)', \dots$$

auf (L. Saalschütz, Vorl., S. 60; siehe auch G. Frobenius, Sitzungsber. d. preuß. Akademie d. W., 1910, S. 826). Durch eine kleine Umformung kann man

daraus auf die höheren Ableitungen von $\frac{1}{1-e^x}$ gelangen. In der Tat ist

$$\left[\frac{1}{1-e^x}\right]' = -\frac{-e^x}{(1-e^x)^2} = \frac{e^x}{(1-e^x)^2},$$

$$\left[\frac{1}{1-e^x}\right]'' = \left[\frac{e^x}{(1-e^x)^2}\right]' = \frac{e^x}{(1-e^x)^2} - 2\frac{e^x \cdot e^x}{(1-e^x)^3} = \frac{e^x + e^{2x}}{(1-e^x)^3},$$

$$\left[\frac{1}{1-e^x}\right]''' = \left[\frac{e^x + e^{2x}}{(1-e^x)^3}\right]' = \frac{e^x + 2e^{2x}}{(1-e^x)^3} - 3\frac{(e^x + e^{2x}) \cdot e^x}{(1-e^x)^4} = \frac{e^x + 4e^{2x} + e^{3x}}{(1-e^x)^4},$$

und wenn

$$\left[\frac{1}{1-e^x}\right]^{(n)} = \frac{C_0^n e^x + C_1^n e^{2x} + \dots + C_{n-1}^n e^{nx}}{(1-e^x)^{n+1}}$$

gesetzt wird, so folgt

$$\begin{aligned} \left[\frac{1}{1-e^x}\right]^{(n+1)} &= \frac{C_0^n e^x + 2C_1^n e^{2x} + 3C_2^n e^{3x} + \dots + nC_{n-1}^n e^{nx}}{(1-e^x)^{n+1}} - \\ &\quad - \frac{n+1}{(1-e^x)^{n+2}} (C_0^n e^x + C_1^n e^{2x} + \dots + C_{n-2}^n e^{(n-1)x} + C_{n-1}^n e^{nx}) \cdot -e^x \\ &= \frac{\begin{cases} C_0^n e^x + 2C_1^n e^{2x} + \dots + nC_{n-1}^n e^{nx} - C_0^n e^{2x} - 2C_1^n e^{3x} - \dots - \\ - (n-1)C_{n-2}^n e^{nx} - nC_{n-1}^n e^{(n+1)x} + (n+1)C_0^n e^{2x} + \\ + (n+1)C_1^n e^{3x} + \dots + (n+1)C_{n-2}^n e^{nx} + (n+1)C_{n-1}^n e^{(n+1)x} \end{cases}}{(1-e^x)^{n+2}} \\ &= \frac{\begin{cases} C_0^n e^x + (2C_1^n + nC_0^n) e^{2x} + (3C_2^n + (n-1)C_1^n) e^{3x} + \dots + \\ + (nC_{n-1}^n + 2C_{n-2}^n) e^{nx} + C_{n-1}^n e^{(n+1)x} \end{cases}}{(1-e^x)^{n+2}} \end{aligned}$$

und nach der Rücklaufformel

$$\left[\frac{1}{1-e^x}\right]^{(n+1)} = \frac{C_0^{n+1} e^x + C_1^{n+1} e^{2x} + C_2^{n+1} e^{3x} + \dots + C_{n-1}^{n+1} e^{nx} + C_n^{n+1} e^{(n+1)x}}{(1-e^x)^{n+2}}$$

womit die allgemeine Gültigkeit der Formel nachgewiesen ist.

Im Zusammenhang damit steht das Auftreten der Zahlen C_k^n in der Entwicklung von $\frac{p-1}{p-c^u}$ nach Potenzen von u (L. Saalschütz, Vorl., S. 74).

Es sei noch auf eine Abhandlung von E. Unferdinger im Archiv der Mathematik und Physik 41 (1864), S. 145–151 hingewiesen.

9. *Zusammenhang der Zahlen C_k^n mit gewissen Reihen.* Hier mögen noch kurz einige Abhandlungen von L. Toscano, u. a. Tohoku math. Journal 38 (1933), S. 332–342; Boll. Un. Mat. Ital. 15 (1936), S. 8–12; 16 (1937), S. 144–149 angeführt werden. In der zweiten sind insbesondere auch die in 6. angegebenen Zahlenwerte enthalten.

(Eingegangen am 23. 6. 1941.)

Modell einer Differentialrechnung mit aktual unendlich kleinen Größen erster Ordnung.

Von

Ludwig Neder in Münster i. W.

Zweck der vorliegenden Arbeit ist der Nachweis der Widerspruchsfreiheit einer durch Weglassung der unendlich kleinen Größen höherer Ordnung vereinfachten, im übrigen aber ganz Leibnizischen Differentialrechnung mit aktual unendlich kleinen Größen (Differentialen) erster Ordnung¹⁾.

Dabei tritt an die Stelle des Satzes, daß das Produkt zweier unendlich kleinen Größen (1. Ordnung) eine unendlich kleine Größe 2. Ordnung ist, der Satz, daß dieses Produkt exakt verschwindet: Es wird also genau so gerechnet, wie in der guten alten Zeit der Differentialrechnung.

Die Grundlage unserer Betrachtungen ist ein System von komplexen Zahlen der Gestalt

$$A = a_1 + a_2 \eta, \text{ wo } a_1, a_2 \text{ reelle Zahlen und } \eta^2 = 0.$$

Dieses System ist nun keineswegs neu. Es kommt vielmehr (als System der dualen Zahlen) in der Literatur vor, neben anderen, allgemeineren Systemen, in denen es als Teilsystem oder als Spezialfall enthalten ist.

Daß solche Systeme für geometrische Zwecke bereits Verwendung gefunden haben, entnehme ich einer liebenswürdigen Mitteilung des Herrn van der Waerden.

Im einzelnen sind zunächst folgende Verfasser anzuführen: A. Kotjelnikoff²⁾, D. Seiliger³⁾, Joh. Petersen⁴⁾, E. Study⁵⁾, J. Grünwald⁶⁾. Sodann erkannte P. Predella⁷⁾, daß das obengenannte komplexe Zahlensystem auch als

¹⁾ Optimisten mögen hoffen, daß man die hier vorgetragenen Betrachtungen eines Tages dazu verwenden wird, eine neue (und keineswegs zu verachtende) Begründung der Differentialrechnung zu geben.

²⁾ Schraubenberechnung und einige Anwendungen derselben auf Geometrie und Mechanik. Kasan Univ. 1895/96.

³⁾ Hauptformeln der komplexen Liniengeometrie. Kasan Univ. Nr. 12 (1897).

⁴⁾ Nouveau principe pour études de géométrie des droites. Oversigt over det K. Danske Vidensk. Selskabs Forhandl. 1898.

⁵⁾ Wir zitieren nur: Geometrie der Dynamen, S. 195ff. Leipzig 1903.

⁶⁾ Über duale Zahlen und ihre Anwendung in der Geometrie. Mon. hefte f. Math. 17 (1906).

⁷⁾ Saggio di Geometria non-Archimedeae. (Nota II.) Giorn. di Mat. di Battaglini 50 (1911). Vgl. Abschn. X.

(„eindimensionales“) nicht-archimedisches Zahlssystem interpretiert werden kann. (Er sagt auch geradezu: $a_2\eta$ heißt unendlich klein.) Weiter sind dann noch zu nennen C. Segre⁸⁾ und Chr. Betsch⁹⁾.

Auch der Funktionsbegriff, den wir benutzen, kommt (bis auf eine Kleinigkeit) in der Literatur bereits vor: Petersen hat — allerdings nur für ganze rationale, exponentielle und goniometrische Funktionen, sowie für deren Umkehrungen — in anderer Bezeichnung die Eigenschaft

$$f(a_1 + a_2\eta) = f(a_1) + f'(a_1) \cdot a_2\eta,$$

die er aber nicht zu einer Definition erhebt. Und Predella definiert

$$f(x_1 + x_2\eta, y_1 + y_2\eta, \dots) = f(x_1, y_1, \dots) + \left(x_2 \frac{\partial f}{\partial x_1} + y_2 \frac{\partial f}{\partial y_1} + \dots\right)\eta,$$

was mit unserer Definition der Funktion beinahe übereinstimmt.

§ 1.

Die Arithmetik des Bereiches der Veränderlichen.

1. Das bereits erwähnte *Größensystem*, in dem unsere Gedankengänge verlaufen werden, besteht aus den geordneten Paaren reeller Zahlen

$$A = [a, \alpha], \quad B = [b, \beta], \quad \dots, \quad X = [x, \xi], \quad \dots$$

Wir bezeichnen sie als *gemischte* Größen und führen weiterhin folgende Bezeichnungen ein:

$[a, \alpha]$ heißt *endlich*, wenn $a \neq 0$;

$[0, \alpha]$ heißt *unendlich klein*;

$[a, 0]$ heißt *real*.

Wir rechnen also $[0, 0]$ zu den unendlich kleinen Größen.

2. Die (*exakte*) Gleichheit ist erklärt durch

$$A = B \Leftrightarrow a = b, \alpha = \beta.$$

Diese Beziehung ist offenbar reflexiv, symmetrisch und transitiv.

3. Die *Ordnung* wird hergestellt durch die hinzutretende Definition¹⁰⁾:

$$A < B \Leftrightarrow a < b \quad \text{oder} \quad a = b, \alpha < \beta.$$

⁸⁾ Le geometrie proiettive nei campi di numeri duali. Atti R. Accad. delle Sc. di Torino 47 (1912).

⁹⁾ Zur analytischen Geometrie der dualen Größen. Diss. Tübingen 1917, gedruckt 1918.

¹⁰⁾ Daher ist das System der gemischten Größen ebenso sehr oder ebenso wenig anschaulich wie das System der reellen Zahlen.

Die so erklärte Kleinerheit ist transitiv. Da ferner die drei Möglichkeiten

$$(a < b \text{ oder } a = b, \alpha < \beta) \text{ oder } (a = b, \alpha = \beta) \\ \text{oder } (a = b, \alpha > \beta \text{ oder } a > b)$$

eine vollständige Disjunktion bilden, ist auch die Disjunktion

$$A < B \text{ oder } A = B \text{ oder } A > B$$

vollständig.

4. Die Addition ist erklärt durch

$$A + B = [a + b, \alpha + \beta]^{11)}.$$

Wegen dieser Definition gehorcht sie offenbar den üblichen Gesetzen (vor allem dem kommutativen und dem assoziativen) und besitzt einen Modul (auch Nullelement genannt)

$$O = [0, 0].$$

Auch das Monotoniegesetz

$$A < A' \rightarrow A + B < A' + B$$

besteht zu Recht, wie leicht zu sehen ist.

Ferner gilt: Die Summe zweier unendlich kleinen Größen ist unendlich klein.

5. Die Subtraktion. Die Differenz $A - B$ ist erklärt als die Auflösung der Gleichung

$$X + B = A.$$

Aus unseren Definitionen folgt

$$x + b = a, \quad \xi + \beta = \alpha$$

und daher

$$A - B = [a - b, \alpha - \beta].$$

Und es gilt: Die Differenz zweier unendlich kleinen Größen ist unendlich klein.

6. Die Fast-Gleichheit. Da man wünscht, zwei gemischte Größen bereits dann zusammenzuwerfen, wenn sie in ihren realen Bestandteilen übereinstimmen, hat man Veranlassung, den folgenden Begriff der Fast-Gleichheit einzuführen:

$$A \div B \rightleftharpoons a = b.$$

¹¹⁾ Daher ist

$$[a, \alpha] = [a, 0] + [0, \alpha],$$

also jede gemischte Größe die Summe einer realen und einer unendlich kleinen Größe.

Es gilt dann: Die Differenz zweier fast gleicher Größen ist unendlich klein. Und man hat die Subsumptionsregel

$$A = B \rightarrow A \div B.$$

7. Die Multiplikation zweier gemischter Größen ist erklärt durch

$$A \cdot B = [a\beta, a\beta + b\alpha].$$

Diese Operation ist offenbar kommutativ und assoziativ. Auch ist sie distributiv zur Addition. Und sie besitzt einen Modul (auch Einselement genannt)

$$E = [1, 0].$$

Ferner gilt: Das Produkt zweier endlichen Größen ist endlich. Das Produkt einer endlichen und einer unendlich kleinen Größe ist unendlich klein. Das Produkt zweier unendlich kleinen Größen ist exakt gleich Null¹²⁾.

Letzteres ist (da wir unendlich kleine Größen höherer Ordnung nicht besitzen) der hier angemessene Ausdruck der Tatsache, daß das Produkt zweier unendlich kleinen Größen (1. Ordnung) gegen jede unendlich kleine Größe (1. Ordnung) zu vernachlässigen ist.

8. Die Division. Der Quotient $A : B = \frac{A}{B}$ ist erklärt als die Auflösung der Gleichung

$$X \cdot B = A.$$

In dieser Hinsicht gilt hier Folgendes:

a) Durch eine endliche Größe kann jede gemischte Größe dividiert werden. Denn man hat infolge der letzten Gleichung:

$$xb = a, \quad x = \frac{a}{b}; \quad \frac{a}{b}\beta + b\xi = a, \quad \xi = \frac{b\alpha - a\beta}{b^2},$$

so daß sich ergibt

$$\frac{A}{B} = \left[\frac{a}{b}, \frac{b\alpha - a\beta}{b^2} \right] \quad \text{bei } b \neq 0.$$

b) Durch eine unendlich kleine Größe kann nur eine unendlich kleine Größe dividiert werden. Denn die Definitionsgleichung des Quotienten ergibt als Folge von $b = 0$ die Beziehung $a = 0$.

c) Der Quotient zweier unendlich kleinen Größen ist nur bestimmt bis auf die willkürlich bleibende zweite Komponente ξ . Denn man hat

$$[x, \xi] \cdot [0, \beta] = [0, x\beta + 0 \cdot \xi] = [0, \alpha], \quad x = \frac{\alpha}{\beta}, \quad \xi \text{ beliebig.}$$

Daher wird hier

$$\frac{A}{B} \div \left[\frac{\alpha}{\beta}, 0 \right] \quad \text{bei } \beta \neq 0 \quad \text{oder } B \neq 0.$$

¹²⁾ Deshalb kann das Monotoniegesetz nicht gelten.

Die üblichen Quotientenformeln gelten für das Rechnen mit unendlich kleinen Größen nur zum Teil. So ist z. B. die Formel

$$A = \frac{A \cdot B}{B}$$

falsch, dagegen die (später gebrauchte) Formel

$$A = \frac{A}{B} \cdot B \quad \text{bei } B \neq 0$$

richtig, weil man hat

$$[0, \alpha] = \left[\frac{\alpha}{\beta}, \xi \right] \cdot [0, \beta],$$

ganz gleichgültig, welchen Wert ξ auch haben mag.

9. *Ein wichtiger Satz:* Die Beziehungen

$$A \doteq B \quad \text{und} \quad A \cdot [0, \gamma] = B \cdot [0, \gamma]$$

sind bei $[0, \gamma] \neq 0$ (d. h. bei $\gamma \neq 0$) für einander notwendig und hinreichend.

Beweis. Die beiden Beziehungen besagen

$$[a, \alpha] \doteq [b, \beta], \quad \text{d. h.} \quad a = b$$

beziehungsweise

$$[a, \alpha] \cdot [0, \gamma] = [b, \beta] \cdot [0, \gamma] \quad \text{oder} \quad [0, a\gamma] = [0, b\gamma],$$

$$\text{d. h.} \quad a\gamma = b\gamma \quad \text{bei } \gamma \neq 0, \quad \text{also } a = b.$$

10. *Das System der realen Größen* $[a, 0]$ gehorcht, wie die vorstehenden Betrachtungen zeigen, den folgenden Gesetzen:

$$[a, 0] = [b, 0] \Leftrightarrow a = b, \quad [a, 0] < [b, 0] \Leftrightarrow a < b,$$

$$[a, 0] \pm [b, 0] = [a \pm b, 0],$$

$$[a, 0] \cdot [b, 0] = [ab, 0], \quad [a, 0] : [b, 0] = \left[\frac{a}{b}, 0 \right] \quad \text{bei } b \neq 0.$$

Diese Gesetze zeigen, daß das genannte System vollkommen isomorph ist mit dem System der reellen Zahlen. Für alle Anwendungen kann daher ersteres an die Stelle des letzteren treten. Und man darf dann die Stellvertreter mit dem Namen der Vertretenen bezeichnen:

$$[a, 0] = a.$$

§ 2.

Die Funktion einer gemischten Veränderlichen.

1. *Der Funktionsbegriff.* Wir haben nicht die Absicht, hier die allgemeine gemischte Funktion einer gemischten Veränderlichen zu studieren. Vielmehr wollen wir unsere Betrachtungen beschränken auf die „Deckfunktion“ $F(X) = F[x, \xi]$, die zu einer reellen Funktion $f(x)$ gehört, und die folgendermaßen erklärt ist:

Definition. Als gemischte Funktion, die eine gegebene reelle, differenzierbare Funktion $f(x)$ „deckt“, bezeichnen wir den Ausdruck

$$F(X) = F[x, \xi] = [f(x), \xi f'(x)].$$

Setzen wir darin speziell $f(x) = c$ (konstant), so erhalten wir

$$C = [c, 0],$$

eine Konstante, die wir auch mit c bezeichnen dürfen und wollen. Zur Deckfunktion gehört also notwendig die (übrigens reale) „Deckkonstante“.

Man kann aber auch anderweitig einsehen, daß unreaie Konstanten unzulässig sind. Nämlich deshalb, weil man mit ihnen eine vorliegende Deckfunktion nicht immer multiplizieren kann, ohne den Funktionsbereich zu verlassen. Denn man hat, wenn C jetzt eine beliebige gemischte Konstante bedeutet:

$$C \cdot F(X) = [c, \gamma] \cdot [f(x), \xi f'(x)] = [cf(x), \xi c f'(x) + \gamma f(x)],$$

und dieser Ausdruck fällt (im wesentlichen) nur für $\gamma = 0$ unter unseren Funktionsbegriff.

Bemerkung. Daß Summe, Differenz, Produkt und Quotient zweier Deckfunktionen wieder eine Deckfunktion ist, kann leicht bewiesen werden, soll aber hier nicht benutzt werden, da man zum Beweise den Ableitungskalkül braucht.

2. *Folgerung aus vorstehender Definition.* Wenn man setzt

$$X_1 = [x, \xi_1], \quad X_2 = [x, \xi_2],$$

so ergibt sich

$$\begin{aligned} F(X_1) - F(X_2) &= [f(x), \xi_1 f'(x)] - [f(x), \xi_2 f'(x)] \\ &= [0, (\xi_1 - \xi_2) f'(x)] = [0, \xi_1 - \xi_2] \cdot [f'(x), 0] = (X_1 - X_2) \cdot [f'(x), 0]. \end{aligned}$$

Nach § 1, Nr. 8, c) ist daher

$$\frac{F(X_1) - F(X_2)}{X_1 - X_2} \doteq [f'(x), 0],$$

was unabhängig ist von X_1 und X_2 .

Dies besagt ganz genau die den älteren Analysten wohlvertraute Tatsache, daß die Funktion in ihren kleinsten Teilen geradlinig verläuft.

3. *Die ganzen rationalen Funktionen*¹³⁾ sind erklärt durch die Formel

$$G(X) = G[x, \xi] = [g(x), \xi g'(x)], \quad \text{wo} \quad g(x) = \sum_{v=0}^n c_v x^v \quad \text{mit} \quad c_n \neq 0.$$

¹³⁾ Hier wird ausnahmsweise die Formel für die Differentiation einer Summe und eines Produktes verwendet.

Dazu tritt noch die Funktion, die identisch verschwindet. Dann gelten folgende Sätze:

a) Jede ganze rationale Funktion besitzt eine *endliche Potenzreihenentwicklung*. Denn man hat, wenn die auftretenden Summen von 0 bis n laufen:

$$\begin{aligned} G(X) &= [\Sigma c, x^r, \xi \Sigma v c, x^{r-1}] \\ &= \Sigma [c, x^r, \xi v c, x^{r-1}] = \Sigma [c, 0] \cdot [x^r, \xi v x^{r-1}] = \Sigma c, X^r. \end{aligned}$$

b) Es gilt ein *Nullitätssatz*: Eine ganze rationale Funktion vom Grade $n \geq 1$ besitzt höchstens n „Wurzelgebilde“. Denn die Annahme

$$G[x_r, \xi_r] = 0 \quad \text{gibt} \quad g(x_r) = 0 \quad \text{und} \quad \xi_r g'(x_r) = 0$$

und umgekehrt. Und man erkennt so:

a) Jeder einfachen Wurzel von $g(x)$ entspricht ein (übrigens realer) *Wurzepunkt* von $G(X)$, nämlich

$$X_r = [x_r, 0].$$

β) Jeder mehrfachen Wurzel von $g(x)$ entspricht eine *Wurzelstrecke* von $G(X)$, gegeben durch

$$X_r = [x_r, \xi], \quad \text{wo} \quad -\infty < \xi < +\infty.$$

c) Daraus folgt der *Identitätssatz*: Wenn eine ganze rationale Funktion vom Grade $\leq n$ in $(n+1)$ Punkten verschwindet, von denen keine zwei fast-gleich sind, so verschwindet sie identisch.

d) Wenn ein realer höchster Koeffizient $c = [c, 0]$ und n reale und verschiedene Wurzelpunkte $X_r = [x_r, 0]$, $r = 1, 2, \dots, n$ gegeben sind, so besteht eine *Produktentwicklung*

$$\begin{aligned} c \cdot \Pi(X - X_r) &= [c, 0] \cdot \Pi[x - x_r, \xi] \\ &= [c, 0] \cdot [\Pi(x - x_r), \xi \Sigma \Pi^{(v)}(x - x_r)] \\ &= [c \Pi(x - x_r), \xi c \Sigma \Pi^{(v)}(x - x_r)] = G(X). \end{aligned}$$

$\Pi^{(v)}$ bedeutet, daß der v -te Faktor wegzulassen ist.

4. Die *goniometrischen Funktionen*. Von diesen sind die beiden ersten erklärt durch die Formeln

$$\sin X = [\sin x, \xi \cos x], \quad \cos X = [\cos x, -\xi \sin x].$$

Indem wir darauf verzichten, die Additionstheoreme zu verifizieren, stellen wir nur fest, daß der Pythagoras gilt:

$$\begin{aligned} \cos^2 X + \sin^2 X &= [\cos^2 x, -2\xi \sin x \cos x] + [\sin^2 x, 2\xi \sin x \cos x] \\ &= [1, 0] = E. \end{aligned}$$

Ferner sieht man sofort, daß der $\sin X$ in den Punkten

$$X_r = [r\pi, 0], \quad r = 0, \pm 1, \pm 2, \dots,$$

verschwindet, und sonst nirgends.

Weiter ist die Funktion $\tan X$ zu erklären durch die Gleichung

$$\tan X = \left[\tan x, \xi \frac{1}{\cos^2 x} \right].$$

Sie genügt der Beziehung

$$\tan X = \sin X : \cos X = \left[\frac{\sin x}{\cos x}, \xi \frac{\cos x \cos x + \sin x \sin x}{\cos^2 x} \right] \quad \text{bei } \cos x \neq 0.$$

5. Die *Exponentialfunktion* ist für jeden Wert von X erklärt durch die Formel

$$e^X = [e^x, \xi e^x].$$

Sie ist stets endlich (also niemals Null), da e^x nirgends verschwindet. Und es gilt das Additionstheorem:

$$\begin{aligned} e^{X_1} \cdot e^{X_2} &= [e^{x_1}, \xi_1 e^{x_1}] \cdot [e^{x_2}, \xi_2 e^{x_2}] \\ &= [e^{x_1} e^{x_2}, e^{x_1} \xi_2 e^{x_2} + \xi_1 e^{x_1} e^{x_2}] = [e^{x_1 + x_2}, (\xi_1 + \xi_2) e^{x_1 + x_2}] \\ &= e^{X_1 + X_2}. \end{aligned}$$

6. Der *natürliche Logarithmus* ist für jeden endlichen „positiven“ Wert von X erklärt als die Auflösung der Gleichung

$$e^Y = X.$$

Diese liefert

$$[e^y, \eta e^y] = [x, \xi], \quad \text{also } e^y = x, y = \log x; \quad \eta x = \xi, \quad \eta = \xi \frac{1}{x}.$$

Als Wert des natürlichen Logarithmus ergibt sich daher

$$\log X = \left[\log x, \xi \frac{1}{x} \right],$$

was ganz genau unter unseren Funktionsbegriff fällt.

Das Multiplikationstheorem folgt aus der Definition der Funktion und dem Additionstheorem der Exponentialfunktion:

$$\begin{aligned} e^{Y_1} &= X_1, \quad e^{Y_2} = X_2; \quad e^{Y_1 + Y_2} = X_1 X_2; \\ \log(X_1 X_2) &= Y_1 + Y_2 = \log X_1 + \log X_2. \end{aligned}$$

7. Die *drei irrationalen Rechenoperationen* können immer dann eingeführt werden, wenn eine Exponentialfunktion und ein natürlicher Logarithmus als gegeben vorliegen. Man hat dann zunächst zu setzen

$$A^B = e^{B \cdot \log A}$$

und findet dann für die beiden Umkehrungen

$$\sqrt[B]{A} = A^{E:B}; \quad \underset{B}{\wedge} A = \frac{\log A}{\log B}.$$

Die Rechengesetze ergeben sich dann (soweit sie gelten) wie üblich.

§ 3.

Die Differentialrechnung der Funktionen einer Veränderlichen.

Es handelt sich hier um Gleichungen, die einerseits zwischen Differentialen und andererseits zwischen Differentialquotienten und gemischten Größen stattfinden. Dabei ist zu beachten, daß bei ersteren nur die exakte Gleichheit sinnvoll ist, während bei letzteren die Fast-Gleichheit in Aktion tritt, da es sich hier um Quotienten von unendlich kleinen Größen handelt, die in der zweiten Komponente unbestimmt bleiben:

1. *Differentiale.* Unter dem Differential der unabhängigen Veränderlichen $X = [x, \xi]$ verstehen wir die unendlich kleine Größe

$$dX = [0, \xi'],$$

worin ξ' willkürlich bleibt.

Unter dem Differential der abhängigen Veränderlichen, genommen im Punkte $X = [x, \xi]$, verstehen wir die (wie sich ergeben wird) gleichfalls unendlich kleine Größe

$$dF(X) = F(X + dX) - F(X).$$

2. *Der Differentiationsprozeß* verläuft nunmehr wie folgt. Es ist zunächst:

$$\begin{aligned} dF(X) &= F[x, \xi + \xi'] - F[x, \xi] \\ &= [f(x), (\xi + \xi') f'(x)] - [f(x), \xi f'(x)] = [0, \xi' f'(x)] \\ &= [f'(x), 0] \cdot [0, \xi']. \end{aligned}$$

Man hat also

$$(1) \quad dF(X) = [f'(x), 0] \cdot dX.$$

Diese Differentialgleichung kann nach § 1, Nr. 8, c) bei $dX \neq 0$ geschrieben werden

$$\frac{dF(X)}{dX} \cdot dX = [f'(x), 0] \cdot dX,$$

so daß sie nach § 1, Nr. 9 (interessanterweise) vollkommen gleichwertig ist mit der Differentialquotienten-Fast-Gleichung

$$\frac{dF(X)}{dX} \doteq [f'(x), 0].$$

3. *Die Grundformeln* der Differentiation ergeben sich jetzt mit einem Schlage wie folgt: Wir setzen zunächst, unter der hier erfüllten Voraussetzung, daß auch $f''(x)$ existiert,

$$F^*(X) = [f'(x), \xi f''(x)],$$

was also nicht die Ableitung der gemischten Funktion $F(X)$ vorstellt, sondern die Deckfunktion der reellen Funktion $f'(x)$. Man konstatiert dann sofort, daß die Beziehung (1) gleichbedeutend ist mit

$$dF(X) = F'(X) \cdot dX.$$

Und genau so wie in Nr. 2 sieht man, daß diese Differentialgleichung gleichwertig ist mit der Differentialquotienten-Fast-Gleichung

$$\frac{dF(X)}{dX} \doteq F'(X).$$

Man hat daher ohne weiteres sämtliche Grundformeln der Differentiation, wie z. B. die Differentialgleichungen

$$dX^n = nX^{n-1} \cdot dX, \quad de^X = e^X \cdot dX, \quad d \sin X = \cos X \cdot dX,$$

oder die Differentialquotienten-Fast-Gleichungen

$$\frac{dX^n}{dX} \doteq nX^{n-1}, \quad \frac{de^X}{dX} \doteq e^X, \quad \frac{d \sin X}{dX} \doteq \cos X.$$

4. *Noch eine Differentiation.* Wenn auf eine gegebene Funktion eine Operation ausgeübt wird, oder mehrere gegebene Funktionen miteinander verknüpft werden, so besteht die Möglichkeit, daß man nicht weiß oder nicht benutzen will, ob in der zweiten Komponente der erhaltenen Funktion die Ableitung der ersten Komponente vorschriftsgemäß auftritt. Man hat dann Veranlassung, die Funktion anzusetzen in der Gestalt

$$F(X) = [f(x), \xi \varphi(x)],$$

und erhält daraus, wenn man die Rechnungen aus Nr. 2 wiederholt,

$$dF(X) = [\varphi(x), 0] \cdot dX \quad \text{oder} \quad \frac{dF(X)}{dX} \doteq [\varphi(x), 0].$$

5. *Die kombinatorischen Formeln der Differentiation.* Man hat zunächst

$$\text{bei } U(X) = [u(x), \xi u'(x)] \quad \text{und} \quad c \cdot U(X) = [cu(x), \xi cu'(x)],$$

und zwar ohne Benutzung der Tatsache, daß $cu'(x)$ die Ableitung von $cu(x)$ ist,

$$\frac{dcU(X)}{dX} \doteq [cu'(x), 0] = c \cdot [u'(x), 0] \doteq c \frac{dU(X)}{dX}.$$

Die weiter noch in Betracht kommenden Formeln

$$\frac{d(U \pm V)}{dX} \doteq \frac{dU}{dX} \pm \frac{dV}{dX},$$

$$\frac{d(U \cdot V)}{dX} \doteq V \cdot \frac{dU}{dX} + U \cdot \frac{dV}{dX},$$

$$\frac{d(U:V)}{dX} \doteq \frac{V \cdot \frac{dU}{dX} - U \cdot \frac{dV}{dX}}{V^2},$$

$$\frac{dF(U(X))}{dX} \doteq \frac{dF(U)}{dU} \cdot \frac{dU(X)}{dX}$$

können auf Grund der vorstehenden Betrachtungen direkt bewiesen werden. Dies ist jedoch nicht nötig, da die Formeln in der im nächsten Paragraphen bewiesenen Differentiationsformel (2) enthalten sind.

§ 4.

Die Differentialrechnung der Funktionen zweier Veränderlichen.

1. Unter *einer Funktion zweier Variablen* $X = [x, \xi]$, $Y = [y, \eta]$ verstehen wir den Ausdruck

$$F(X, Y) = F([x, \xi], [y, \eta]) = [f(x, y), \xi f_1(x, y) + \eta f_2(x, y)].$$

Wir betrachten also auch hier nur die Funktion, die eine gegebene reelle Funktion $f(x, y)$ „deckt“.

Nebenbei stellen wir fest, daß jede Funktion einer gemischten Veränderlichen X oder Y auch eine Funktion zweier gemischter Variablen X und Y ist.

2. Unter *den Differentialen der unabhängigen Veränderlichen* X bzw. Y verstehen wir die unendlich kleinen Größen

$$dX = [0, \xi'] \quad \text{bzw.} \quad dY = [0, \eta'].$$

3. Unter *den partiellen Differentialen der abhängigen Veränderlichen* verstehen wir die gleichfalls unendlich kleinen Größen

$$\begin{aligned} \partial_1 F &= F(X + dX, Y) - F(X, Y) \\ &= F([x, \xi + \xi'], [y, \eta]) - F([x, \xi], [y, \eta]) = [0, \xi' f_1(x, y)], \\ \partial_2 F &= F(X, Y + dY) - F(X, Y) \\ &= F([x, \xi], [y, \eta + \eta']) - F([x, \xi], [y, \eta]) = [0, \eta' f_2(x, y)]. \end{aligned}$$

Nach § 3 gilt dann

$$\begin{aligned} \frac{\partial_1 F}{dX} &\doteq [f_1(x, y), 0], \\ \frac{\partial_2 F}{dY} &\doteq [f_2(x, y), 0]. \end{aligned}$$

Wie man sieht, müssen im Zähler runde, im Nenner aber gerade d stehen.

Im Sinne des vorliegenden Aufbaues der Differentialquotienten sind also die runden ∂ im Nenner nicht zu vertreten.

4. Ferner *gilt der Satz*: Wenn

$$F(X, Y) = [f(x, y), \xi \varphi(x, y) + \eta \psi(x, y)],$$

so gelten die Formeln

$$\begin{aligned} \frac{\partial_1 F}{dX} &\doteq [\varphi(x, y), 0], \\ \frac{\partial_2 F}{dY} &\doteq [\psi(x, y), 0]. \end{aligned}$$

5. Das *vollständige Differential* einer Funktion $F(X, Y)$ im Punkte $X = [x, \xi]$, $Y = [y, \eta]$ ist gegeben durch die Formel

$$dF = F(X + dX, Y + dY) - F(X, Y) = [0, \xi' f_1(x, y) + \eta' f_2(x, y)].$$

Es gilt also die Beziehung

$$dF = \partial_1 F + \partial_2 F.$$

Dieselbe geht nach § 1, Nr. 8, c) bei „Erweiterung“ mit dX bzw. dY über in

$$dF = \frac{\partial_1 F}{dX} \cdot dX + \frac{\partial_2 F}{dY} \cdot dY,$$

eine Formel, die ihren geheimnisvollen Nimbus jetzt vollkommen verloren hat.

6. *Differentiation zusammengesetzter Funktionen.* Ist nunmehr

$$Z = F(U, V), \quad \text{wo} \quad U = U(T), \quad V = V(T),$$

so ist nach 5

$$dF = \frac{\partial_1 F}{dU} \cdot dU + \frac{\partial_2 F}{dV} \cdot dV,$$

und hieraus folgt durch Division mit dT

$$(2) \quad \frac{dF}{dT} = \frac{\partial_1 F}{dU} \cdot \frac{dU}{dT} + \frac{\partial_2 F}{dV} \cdot \frac{dV}{dT},$$

was eine wohlbekannte Formel vorstellt.

Zusatz bei der Korrektur: Formel (2) und ihr Beweis erfordert die Bedingungen

$$dU = dU(T) \neq 0, \quad dV = dV(T) \neq 0.$$

Es ist daher besser, die kombinierenden Formeln der Differentiation (§ 3, Nr. 5) direkt und ohne Heranziehung von (2) zu beweisen, wozu obenstehend alles bereit liegt.

(Eingegangen am 1. 12. 1940.)

Bemerkungen zu einigen von Herrn Collatz angegebenen Eigenwert- abschätzungen bei linearen Integralgleichungen.

Von
Rudolf Iglisch in Braunschweig.

§ 1.

Übersicht.

In einer kürzlich erschienenen Arbeit¹⁾ hat Herr Collatz folgende Fragestellung angeschnitten: Zugrunde gelegt sei die Integralgleichung

$$(1) \quad \varphi(s) - \lambda \int K(s, t) \varphi(t) dt = 0$$

mit *symmetrischem* Kern, für den ein iterierter Kern $K_n(s, t)$ von endlicher Ordnung $n \geq 1$ stetig sei²⁾. Die Buchstaben s und t können mehrdimensionale Punktvariable vertreten, das Integral sei wie alle folgenden über ein und dasselbe endliche Grundgebiet \mathfrak{B} erstreckt. Von einer beliebigen Funktion $F_0(s)$ ausgehend wird

$$(2) \quad F_1(s) = \int K(s, t) F_0(t) dt,$$

allgemein rekursiv

$$(3) \quad F_n(s) = \int K(s, t) F_{n-1}(t) dt$$

gebildet.

Jetzt wird abkürzend bezeichnet

$$(4) \quad a_n = \int F_m(s) F_{n-m}(s) ds \quad (\text{für } n = 0, 1, \dots; m \leq n)$$

und

$$(5) \quad \mu_n^{(0)} = \frac{a_{n-1}}{a_n},$$

weiter rekursiv

$$(6) \quad \mu_n^{(l)} = \frac{\mu_n^{(l-1)} - \lambda_{l+r-1}}{\mu_{n+1}^{(l-1)} - \lambda_{l+r-1}} \mu_{n+1}^{(l-1)};$$

¹⁾ L. Collatz, Schrittweise Näherungen bei Integralgleichungen und Eigenwertschranken. Math. Zeitschr. 46 (1940), S. 692–708. Diese Arbeit werde ich bei künftigen Hinweisen stets kurz mit C. bezeichnen. In meinen Bezeichnungen und Begriffsbildungen werde ich mich an diese Arbeit anschließen.

²⁾ Diese Aussage enthält für $n = 2$ die Voraussetzung der mittleren Stetigkeit bei C.

dabei bedeutet λ , den r -ten Eigenwert (ein mehrfacher Eigenwert wird so oft hingeschrieben, wie seine Vielfachheit angibt) des zunächst als positiv definit angenommenen Kernes $K(s, t)$. Dann ist das grundlegende Ergebnis von C.:

$$(7) \quad \mu_n^{(l)} \geq \mu_{n+1}^{(l)} \geq \lambda_{r+l} \quad (n \geq 3; l = 0, 1, 2, \dots),$$

dabei bedeutet r den niedrigsten Index der bei Entwicklung von $F_1(s)$ nach den Eigenfunktionen $\varphi_r(s)$ von $K(s, t)$ auftretenden Eigenfunktionen; mit anderen Worten: $F_1(s)$ ist zu $\varphi_1(s), \varphi_2(s), \dots, \varphi_{r-1}(s)$ orthogonal, nicht aber zu $\varphi_r(s)$. Beim Beweise dieser Ungleichung (7), aus der die übrigen Ergebnisse der Arbeit C. folgen (mit Ausschluß des Kap. 11, von dem später die Rede sein wird), benutzt Herr Collatz den Entwicklungssatz für $F_1(s)$. In der Einleitung seiner Arbeit weist Herr Collatz selbst darauf hin, daß es wünschenswert wäre, den Entwicklungssatz aus dem Beweise zu entfernen, da man dann statt von $F_1(s)$ gleich von $F_0(s)$ ausgehen könnte, also sich einen Iterationsschritt erspart hätte. Diese Abänderung des Beweisganges von Herrn Collatz bringe ich in § 3. Die Aussage (7) hat dann also für $n \geq 1$ Gültigkeit. Der Beweis von Herrn Collatz und damit auch meine kleine Änderung beruhen auf einem Hilfssatz (C. Kap. 4), zu dessen Beweis Herr Collatz auch den Entwicklungssatz heranzieht. Da dies gleichfalls nicht notwendig ist, stelle ich den Beweis dieses Hilfssatzes in § 2 noch einmal — mit einer unwesentlichen Verschärfung, die auch für die Folgerungen aus dem Hilfssatz zutrifft — voran.

In C. Kap. 11 zeigt Herr Collatz, daß unter sehr speziellen Voraussetzungen zwischen Größtwert und Kleinstwert der Funktion

$$(8) \quad G(s) = \frac{F_0(s)}{F_1(s)}$$

ein Eigenwert λ_k von (1) liegt. In einer brieflichen Mitteilung an den Verfasser teilte Herr Collatz mit, er vermute, daß dieser Satz wesentlich allgemeinere Gültigkeit besitze. Dies ist tatsächlich der Fall, wie in § 5 bewiesen wird²⁾. Der Beweis beruht auf einem Satz über die Änderung der Eigenwerte bei einem Kern mit Parameter, den ich in § 4 voranstelle. Dieser Satz leistet übrigens, wie vielleicht bei anderer Gelegenheit mitgeteilt werden soll, bei vielen anderen Problemen einfachere und bessere Dienste als das bisher zu deren Behandlung herangezogene Minimum-Maximum-Prinzip der Eigenwerte. — In § 7 beschäftige ich mich noch insbesondere mit dem leichter zu behandelnden Fall, daß $K(s, t)$ eine positive Eigenfunktion besitzt. Hier ist für die Anwendungen ein Satz besonders nützlich, der bei positiven Kernen mit im Inneren des Integrationsgebietes wesentlich positiver Spur die Existenz

²⁾ Inzwischen teilte mir Herr Collatz einen anders gearteten Beweis des für die Praxis wichtigsten Teiles dieses Satzes mit. Dieser wird in Kürze veröffentlicht werden.

einer positiven Eigenfunktion sicherstellt; und zwar gehört diese zum absolut kleinsten Eigenwert λ_1 , der zudem einfach und positiv ist. Dieser Satz ist zwar in der Literatur längst bekannt⁴⁾, ich benutze jedoch die Gelegenheit, in § 6 einen mir besonders einfach erscheinenden Beweis hierfür mitzuteilen.

§ 2.

Der Hilfssatz von Collatz.

Man gehe bei Bildung der Größen a_n und μ_n in (4) und (5) statt von der Funktion $F_0(s)$ aus von der abgeänderten Funktion

$$(9) \quad F_0^*(s) = F_0(s) - \sum_{i=1}^r \varphi_i(s) \int F_0(t) \varphi_i(t) dt \quad (r \geq 1),$$

die als nicht identisch verschwindend angenommen sei, und bezeichne die entsprechenden Größen mit a_n^* bzw. μ_n^* . Ist nun $K(s, t)$ ein positiv definiter symmetrischer Kern, so gilt

$$(10) \quad \mu_n^* \geq \lambda_{r+1}.$$

Beweis. Der abgeänderte symmetrische Kern

$$(11) \quad K^*(s, t) = K(s, t) - \sum_{v=1}^r \frac{\varphi_v(s) \varphi_v(t)}{\lambda_v}$$

besitzt nicht mehr die Eigenfunktionen $\varphi_\varrho(s)$ für $\varrho = 1, 2, \dots, r$, alle übrigen Eigenfunktionen $\varphi_\sigma(s)$ mit $\sigma > r$ von $K(s, t)$ sind noch Eigenfunktionen von $K^*(s, t)$ zu den gleichen Eigenwerten, weitere Eigenfunktionen besitzt $K^*(s, t)$ nicht. $K^*(s, t)$ ist mithin wieder positiv definit. Man setze

$$(12) \quad f(s) = F_0^*(s) - \mu_1^* F_1^*(s)$$

und betrachte die Lösungen $u(s, \lambda)$ der inhomogenen Integralgleichung [C. (4. 4)]

$$(13) \quad u(s, \lambda) - \lambda \int K^*(s, t) u(t, \lambda) dt = f(s).$$

Analog C. S. 696 betrachte man die Hilfsfunktion

$$(14) \quad W^*(\lambda) = \int u(s, \lambda) f(s) ds.$$

1. Fall. $f(s) \equiv 0$. Dann ist $F_0^*(s)$ Eigenfunktion von $K(s, t)$ und $K^*(s, t)$ zum Eigenwert μ_1^* , mithin von selbst $\mu_1^* \geq \lambda_{r+1}$, da $F_0^*(s)$ zu $\varphi_\varrho(s)$ ($\varrho = 1, \dots, r$) orthogonal ist.

⁴⁾ Vgl. z. B. O. D. Kellogg, Orthogonal functions sets arising from integral equations. Amer. Journ. of Math. 40 (1918), S. 145–154. Die kurzen Andeutungen in C. S. 707, Absatz 2 sind kein Beweis dieses Satzes; sie sind vielmehr Folgerungen aus ihm. Vgl. dazu etwa O. Blumenthal, Über die Knickung eines Balkens durch Längskräfte. Zeitschr. f. angew. Math. u. Mech. 17 (1937), S. 232–244.

2. Fall. $f(s) \neq 0$. Wir nehmen entgegen unserer Behauptung an, $K^*(s, t)$ besitze zwischen 0 und μ_1^* keinen Eigenwert. Dann wäre $W^*(\lambda)$ nach (14) für $0 \leq \lambda \leq \mu_1^*$ eindeutig bestimmbar. Überdies wäre

$$W^*(0) = \int f^2(s) ds > 0$$

und (vgl. C., S. 696)

$$\frac{dW^*(\lambda)}{d\lambda} = \iint u(s, \lambda) K(s, t) u(t, \lambda) ds dt \geq 0,$$

da $K(s, t)$ positiv definit sein sollte. Aus diesen beiden Tatsachen würde $W^*(\mu_1^*) > 0$ folgen. Andererseits ergibt sich aber, da für $\lambda = \mu_1^*$ (12) und (13) identisch sind, wegen $u(s, \mu_1^*) = F_0^*(s)$

$$W^*(\mu_1^*) = \int F_0^*(s) (F_0^*(s) - \mu_1^* F_1^*(s)) ds = a_0^* - \mu_1^* a_1^* = 0$$

zufolge der zu (5) analogen gesterntten Beziehung. Dieser Widerspruch löst sich nur, wenn der erste Eigenwert von $K^*(s, t)$, also λ_{r+1} , nicht hinter μ_1^* liegt, womit die Behauptung (10) für $n = 1$ bewiesen ist. Für allgemeines n ergibt sich ihre Richtigkeit entsprechend, da mit $F_0^*(s)$ auch alle weiteren $F_n^*(s)$ zu den ersten r Eigenfunktionen orthogonal sind. (Vgl. C. (4. 3)).

§ 3.

Die Hauptungleichung von Collatz.

In diesem Paragraphen soll die fundamentale Aussage (7) für $n = 1, 2, \dots$ bewiesen werden. Dabei wird $K(s, t)$ wieder als positiv definit, symmetrischer Kern vorausgesetzt; über die Verschiedenheit der Eigenwerte braucht keine Annahme gemacht zu werden, wie dies in C. 6 noch zusätzlich geschieht. $F_0(s)$ sei zu $\varphi_\rho(s)$ für $\rho = 1, 2, \dots, r-1$ orthogonal, nicht aber zu $\varphi_r(s)$. Wir schreiben

$$(15) \quad F_0(s) = c_r \varphi_r(s) + f_{r+1}(s) \quad \text{mit} \quad \int f_{r+1}(s) \varphi_r(s) ds = 0.$$

Dann wird [vgl. C. (2. 2)]

$$(16) \quad F_n(s) = \frac{c_r \varphi_r(s)}{\lambda_r^n} + \int K_n(s, t) f_{r+1}(t) dt = \frac{c_r \varphi_r(s)}{\lambda_r^n} + \int K_n^*(s, t) f_{r+1}(t) dt,$$

wo $K_n^*(s, t)$ den n -ten iterierten Kern zum Kern (11) bezeichnet. Wir benutzen jetzt die quadratische Integralform

$$(17) \quad J(K_n^*; F_0) = \iint F_0(s) K_n^*(s, t) F_0(t) ds dt.$$

Mit dieser Schreibweise wird nach (4)

$$(18) \quad a_n = \int F_0(s) F_n(s) ds = J(K_n^*; F_0),$$

$$(19) \quad \mu_n = \frac{a_{n-1}}{a_n} = \frac{J(K_{n-1}^*; F_0)}{J(K_n^*; F_0)},$$

$$(20) \quad \mu_n - \lambda_r = \frac{a_{n-1} - \lambda_r a_n}{a_n} = \frac{J(K_{n-1}^* - \lambda_r K_n^*; F_0)}{J(K_n^*; F_0)} \geq 0.$$

Nach C. 3 findet man

$$(21) \quad \mu_n - \mu_{n+1} \geq 0.$$

Der in (20) auftretende Kern

$$(22) \quad L_{n-1}^{(1)}(s, t) = K_{n-1}^*(s, t) - \lambda_r K_n^*(s, t)$$

ist ein positiv definiter symmetrischer Kern mit den Eigenfunktionen $\varphi_\varrho(s)$ für $\varrho = r+1, r+2, \dots$ zu den Eigenwerten

$$\lambda_\varrho^{n-1}; \left(1 - \frac{\lambda_r}{\lambda_\varrho}\right).$$

Wir betrachten jetzt die Größen

$$(23) \quad \mu_n^{(1)} = \frac{\mu_n - \lambda_r}{\mu_{n+1} - \lambda_r} \mu_{n+1} = \frac{J(L_{n-1}^{(1)}; F_0)}{J(L_n^{(1)}; F_0)}.$$

Diese Formel hat den gleichen Bau wie (19). Daher erhält man wie vorhin

$$(24) \quad \mu_n^{(1)} - \lambda_{r+1} = \frac{J(L_{n-1}^{(2)}; F_0)}{J(L_n^{(1)}; F_0)} \geq 0$$

mit der Abkürzung

$$(25) \quad L_{n-1}^{(2)}(s, t) = L_{n-1}^{(1)}(s, t) - \lambda_{r+1} L_n^{(1)}(s, t).$$

Von diesem Kern zeigt man leicht, daß er $\varphi_{r+1}(s)$ nicht mehr als Eigenfunktion besitzt, aber noch $\varphi_\varrho(s)$ mit $\varrho = r+2, r+3, \dots$ zu den Eigenwerten

$$\lambda_\varrho^{n-1}; \left\{ \left(1 - \frac{\lambda_r}{\lambda_\varrho}\right) \left(1 - \frac{\lambda_{r+1}}{\lambda_\varrho}\right) \right\}.$$

Genau wie vorhin (21) folgt jetzt für die analogen Bildungen

$$(26) \quad \mu_n^{(1)} - \mu_{n+1}^{(1)} \geq 0.$$

Entsprechend C. (6. 7) setzen wir allgemein

$$(27) \quad \mu_n^{(l)} := \frac{\mu_n^{(l-1)} - \lambda_{r+l-1}}{\mu_{n-1}^{(l-1)} - \lambda_{r+l-1}} \mu_{n+1}^{(l-1)} = \frac{J(L_{n-1}^{(l)}; F_0)}{J(L_n^{(l)}; F_0)}$$

mit

$$(28) \quad L_{n-1}^{(l)}(s, t) = L_{n-1}^{(l-1)}(s, t) - \lambda_{r+l-1} L_n^{(l-1)}(s, t).$$

Wieder folgt wie früher

$$(29) \quad \mu_n^{(l)} - \mu_{n+1}^{(l)} \geq 0; \quad \mu_n^{(l)} - \lambda_{r+l} \geq 0.$$

Die $\mu_n^{(l)}$ bilden also für jedes l eine monoton abnehmende, durch λ_{r+l} beschränkte Zahlenfolge, die mithin konvergiert. Da $L_{n-1}^{(l)}(s, t)$ gerade die Eigenfunktionen $\varphi_\varrho(s)$ mit $\varrho \geq r+l$ zu den Eigenwerten

$$\lambda_\varrho^{n-1}; \left\{ \left(1 - \frac{\lambda_r}{\lambda_\varrho}\right) \left(1 - \frac{\lambda_{r+1}}{\lambda_\varrho}\right) \dots \left(1 - \frac{\lambda_{r+l-1}}{\lambda_\varrho}\right) \right\}$$

besitzt, ist zufolge (27)

$$(30) \quad \lim_{n \rightarrow \infty} \mu_n^{(l)} = \lambda_{r+l+q}, \quad \mu_n^{(l)} \geq \mu_{n+1}^{(l)},$$

wo die ganze Zahl $q \geq 0$ folgende Bedeutung hat: $F_0(s)$ ist zu allen $\varphi_\sigma(s)$ mit $\sigma = r+l, r+l+1, \dots, r+l+q-1$ orthogonal, nicht aber zu $\varphi_{r+l+q}(s)$. Die Formeln (29) sind mit unserer Behauptung (7) gleichwertig, Formel (30), die sich auch in C. 6 findet, ist eine genauere Aussage. Hieraus wird am Schluß von C. 6 folgende spezielle Folgerung gezogen: Ist $F_0(s)$ nicht zu $\varphi_{r+l}(s)$ orthogonal, ist $\lambda_{r+l+1} > \lambda_{r+l}$ und n so groß gewählt, daß $\lambda_{r+l+1} > \mu_{n+p}^{(l)}$ ($p = 1, 2, \dots$) ausfällt, so erhält man für λ_{r+l} die Schranken (6. 9).

$$(31) \quad \mu_{n+1}^{(l)} - \frac{\mu_n^{(l)} - \mu_{n+1}^{(l)}}{\frac{\lambda_{r+l+1} - 1}{\mu_{n+1}^{(l)}}} \leq \lambda_{r+l} \leq \mu_{n+1}^{(l)}.$$

§ 4.

Über die Änderung der Eigenwerte bei einem Kern, der in spezieller Weise von einem Parameter abhängt.

Ich beweise hier folgenden

Hauptsatz: $K(s, t)$ sei ein symmetrischer Kern, für den ein iterierter Kern $K_n(s, t)$ von endlicher Ordnung $n \geq 1$ stetig sei, $H(s, \tau)$ sei in beiden Veränderlichen stetig und nach τ differenzierbar. Es sei mit nur reellen Größen

$$(32) \quad \varphi_{r,\tau}(s) - \lambda_{r,\tau} H(s, \tau) \int K(s, t) \varphi_{r,\tau}(t) dt = 0.$$

Dann gilt, wenn durch herangesetzte Striche Ableitungen nach dem Parameter τ bezeichnet werden,

$$(33) \quad (\lambda_{r,\tau} H(s, \tau))' \text{ muß als Funktion von } s \text{ das Zeichen wechseln.}$$

Bemerkung. Daß (32) nur reelle Eigenwerte besitzt, kann etwa so⁵⁾ gezeigt werden: Sei $\Phi(s) + i\Psi(s)$ eine komplexe Eigenfunktion zum Eigenwert λ . Multipliziert man (32) mit $K(z, s)$ und integriert über s , so kommt mit der Abkürzung

$$K_2(z, t) = \int K(z, s) H(s, \tau) K(s, t) ds,$$

$$\int K(z, s) [\Phi(s) + i\Psi(s)] ds = \lambda_{r,\tau} \int K_2(z, t) [\Phi(t) + i\Psi(t)] dt.$$

⁵⁾ Im Anschluß an Joseph Marty, Existence de solutions singulières pour certaines équations de Fredholm. C. R. 150 (1910), S. 1031–1033.

Multiplikation mit $\Phi(z) - i\Psi(z)$ und Integration liefert

$$(34) \quad \iint K(z, s) [\Phi(s) + i\Psi(s)] [\Phi(z) - i\Psi(z)] ds dz \\ = \lambda_{r\tau} \iint K_2(z, t) [\Phi(t) + i\Psi(t)] [\Phi(z) - i\Psi(z)] dz dt.$$

Da infolge der Symmetrie von $K(z, s)$ und $K_2(z, t)$ beide Integrale reell sind, muß nach (34) auch $\lambda_{r\tau}$ reell sein.

Beweis des Hauptsatzes. Differenziation von (32) nach τ liefert

$$(35) \quad \varphi'_{r\tau}(s) - \lambda_{r\tau} H(s, \tau) \int K(s, t) \varphi'_{r\tau}(t) dt \\ = [\lambda_{r\tau} H(s, \tau)]' \int K(s, t) \varphi_{r\tau}(t) dt.$$

Die zu (32) adjungierte Gleichung

$$(36) \quad \varphi_{r\tau}(s) - \lambda_{r\tau} \int K(s, t) H(t, \tau) \varphi_{r\tau}(t) dt = 0$$

besitzt die Eigenfunktion

$$(37) \quad \varphi_{r\tau}^{(s)} = \frac{\varphi_{r\tau}(s)}{H(s, \tau)}.$$

Nun folgt aus (35)

$$\iint [\lambda_{r\tau} H(s, \tau)]' \varphi_{r\tau}(s) K(s, t) \varphi_{r\tau}(t) ds dt = 0$$

und unter Verwendung von (32) und (37)

$$(38) \quad \int [\lambda_{r\tau} H(s, \tau)]' \varphi_{r\tau}^2(s) ds = 0.$$

Hieraus folgt unmittelbar die Behauptung (33).

1. Folgerung. Ist speziell mit einer Konstanten C und nicht konstantem $G(s)$

$$(39) \quad H(s; \tau) = C \cdot \tau + (1 - \tau) G(s) = G(s) + \tau(C - G(s)),$$

so wird

$$(40) \quad [\lambda_{r\tau} H(s, \tau)]' = \lambda'_{r\tau} [G(s) + \tau(C - G(s))] + \lambda_{r\tau}(C - G(s)).$$

Gilt nun

$$(41) \quad G(s) + \tau(C - G(s)) \geq 0 \quad \text{für } 0 \leq \tau \leq 1,$$

so erhält man im Falle

$$(42) \quad C - G(s) \geq 0 \quad \text{sign } \lambda'_{r\tau} = - \text{sign } \lambda_{r\tau},$$

dagegen im Falle

$$(43) \quad C - G(s) \leq 0 \quad \text{sign } \lambda'_{r\tau} = \text{sign } \lambda_{r\tau}.$$

2. Folgerung. Wechselt $C - G(s)$ nicht das Vorzeichen, so muß nach (40) stets $\lambda'_{r\tau} \neq 0$ sein, kann also nicht sein Vorzeichen wechseln. Ersetzt man (41) durch

$$(44) \quad C > 0,$$

so bleiben die Aussagen (42) und (43) erhalten, da für $\tau = 1$ (41) zu Recht besteht.

§ 5.

Der allgemeine Einschließungssatz.

Behauptung. Zwischen Größtwert und Kleinstwert der in (8) und (2) definierten Funktion $G(s)$ liegt mindestens ein Eigenwert des symmetrischen Kernes $K(s, t)$. Dabei sei der Einfachheit halber etwa vorausgesetzt, daß Größtwert und Kleinstwert von $G(s)$ endlich sind.

Beweis. Der Beweis wird ersichtlich erbracht sein, wenn gezeigt ist, daß folgende Aussagen unmöglich sind:

a) $G(s)$ liegt ganz zwischen zwei aufeinanderfolgenden verschiedenen positiven Eigenwerten λ_r^+ und λ_{r+1}^+ oder zwei aufeinanderfolgenden negativen Eigenwerten λ_n^- und λ_{n+1}^- .

b) $G(s)$ liegt zwischen dem ersten negativen Eigenwert λ_1^- und dem ersten positiven Eigenwert λ_1^+ .

c) $G(s)$ liegt oberhalb des größten positiven Eigenwertes λ_n^+ , bzw. bei Fehlen positiver Eigenwerte oberhalb des kleinsten negativen Eigenwertes λ_1^- . Analog: $G(s)$ liegt unterhalb des kleinsten negativen Eigenwertes λ_m^- oder im Falle des Fehlens negativer Eigenwerte ganz unterhalb λ_1^+ .

Ist übrigens $G(s) = \lambda = \text{const.}$, so ist aus (2) und (8) sofort abzulesen, daß $F_0(s)$ Eigenfunktion von $K(s, t)$ zum Eigenwert λ ist. Künftighin braucht also nur der Fall eines nichtkonstanten $G(s)$ berücksichtigt zu werden.

Ich zeige zunächst den Widerspruch der Aussage

a) Wäre z. B.

$$(45) \quad \lambda_r^+ \leq G(s) \leq \lambda_{r+1}^+,$$

so setze man

$$(46) \quad H(s, \tau) = \lambda_r^+ \tau + (1 - \tau) G(s).$$

Für $\tau = 0$ erscheint aus (32)

$$(47) \quad \varphi_{r0}(s) - \lambda_{r0} G(s) \int K(s, t) \varphi_{r0}(t) dt = 0.$$

Diese Gleichung hat nach (2) und (8) die Eigenfunktion $F_0(s)$ zum Eigenwert 1.

Für $\tau = 1$ kommt

$$(48) \quad \varphi_{r1}(s) - \lambda_{r1} \lambda_r^+ \int K(s, t) \varphi_{r1}(t) dt = 0.$$

Da (41) und (43) erfüllt sind, liegt nach der 1. Folgerung aus unserem Hauptsatz in § 4 der r -te positive Eigenwert λ_{r0}^+ von (47) vor dem r -ten positiven Eigenwert λ_{r1}^+ von (48), d. h.

$$(49) \quad \lambda_{r0}^+ < 1.$$

Verwendet man statt (46) die Funktion

$$(50) \quad H^*(s, \tau) = \lambda_{r+1}^+ (1 - \tau) + \tau G(s),$$

so folgt analog

$$(51) \quad \lambda_{r+1,0}^+ > 1.$$

Nach (49) und (51) könnte 1 nicht Eigenwert von (47) sein, was unseren obigen Feststellungen zuwiderläuft⁶⁾.

Genau so zeigt man, daß $G(s)$ nicht zwischen λ_n^- und λ_{n+1}^- liegen kann. Man braucht zu diesem Zwecke übrigens nur von $K(s, t)$ zu $-K(s, t)$ überzugehen.

b) Die Annahme

$$\lambda_1^- \leq G(s) \leq \lambda_1^+$$

wird in gleicher Weise zum Widerspruch geführt, wenn man statt der 1. Folgerung aus § 4 die 2. Folgerung benutzt.

c) Der gleiche Schluß liefert sofort, daß bei Fehlen negativer Eigenwerte nicht

$$G(s) \leq \lambda_1^+ = \lambda_1$$

sein kann.

Macht man die Annahme, daß λ_n^+ der größte positive Eigenwert ist und

$$(52) \quad \lambda_n^+ \leq G(s) \leq M$$

gilt, so folgt wie oben, daß der n -te positive Eigenwert von (47) kleiner ist als 1; 1 kann auch nicht $(n+1)$ -ter positiver Eigenwert von (47) sein, da dieser $(n+1)$ -te positive Eigenwert größer sein müßte als der entsprechende Eigenwert von

$$\varphi(s) - \lambda M \int K(s, t) \varphi(t) dt = 0.$$

Diese Gleichung besitzt aber keinen $(n+1)$ -ten positiven Eigenwert. Durch Übergang zu $-K(s, t)$ folgt die Unmöglichkeit der übrigen unter c) angeführten Aussagen.

§ 6.

Über Integralgleichungen mit positivem Kern.

Es handelt sich hier um einen Beweis für den folgenden wohl auf O. D. Kellogg⁷⁾ zurückgehenden

Satz. *Gilt für den reellen, stetigen, symmetrischen Kern $K(s, t)$*

$$(53) \quad K(s, t) \geq 0$$

und außerdem

$$(54) \quad K(s, s) > 0$$

⁶⁾ Statt dieses indirekten Beweises könnte man auf die gleiche Weise direkt zeigen, daß zwischen Minimum und Maximum von $G(s)$ mindestens ein Eigenwert von $K(s, t)$ liegt, wie dies Herr Collatz in der in Fußnote 3 erwähnten Arbeit tut.

⁷⁾ Vgl. Fußnote 4.

im Innern des Integrationsgebietes \mathfrak{B} , so ist der absolut kleinste Eigenwert λ_1 positiv und einfach, die zugehörige Eigenfunktion $\varphi_1(s)$ ist im Innern von \mathfrak{B} positiv, $-\lambda_1$ ist kein Eigenwert.

Den Beweis führe ich in mehreren Schritten:

I. Bei jedem λ zwischen 0 und dem ersten positiven Eigenwert λ_1 ($0 \leq \lambda < \lambda_1$) hat für eine stetige positive rechte Seite $p(s) > 0$ die Gleichung

$$(55) \quad y(s) - \lambda \int K(s, t) y(t) dt = p(s)$$

eine positive Lösung $y(s) > 0$.

Beweis. Für $\lambda = 0$ ist der Satz richtig. Man lasse λ bis λ_1 wachsen. $y(s)$ ändert sich stetig mit λ . Würde zum ersten Male ein Wert $\lambda_0 < \lambda_1$ existieren, für den $y(s_0) = 0$ ist, so liefert die Integralgleichung (55) für $s = s_0$ einen Widerspruch.

II. Ist nun λ_1 einfacher Eigenwert, so kann die zugehörige Eigenfunktion $\varphi_1(s)$ nicht das Vorzeichen wechseln, sie darf also als nicht negativ angenommen werden. Andernfalls gäbe es nämlich eine zu $\varphi_1(s)$ orthogonale positive Funktion $p(s)$ derart, daß Gleichung (55) für $\lambda = \lambda_1$ lösbar ist, und zwar kann man durch Zufügen eines passenden Vielfachen von $\varphi_1(s)$ erreichen, daß die Lösung $y(s)$ Werte beiderlei Vorzeichens annimmt. Man wähle ein genügend kleines positives ε und betrachte die Gleichung

$$(56) \quad y(s) - (\lambda_1 - \varepsilon) \int K(s, t) y(t) dt = p(s) + \varepsilon \int K(s, t) y(t) dt.$$

Diese hat bei immer noch positiver rechter Seite eine durch Null gehende Lösung $y(s)$ für $\lambda_1 - \varepsilon < \lambda_1$, was I widerspricht.

III. Die Bedingung (54) sorgt dann dafür, daß $\varphi_1(s)$ im Innern von \mathfrak{B} wirklich positiv ist. Sei etwa s_1 eine isolierte Nullstelle im Innern, so liefert

$$(57) \quad \varphi_1(s_1) - \lambda_1 \int K(s_1, t) \varphi_1(t) dt = 0$$

sofort einen Widerspruch. Die Annahme ganzer Bereiche von Nullstellen wird ebenso zum Widerspruch geführt, wenn s_1 als Randpunkt eines solchen Nullstellenbereiches gewählt wird.

IV. Unter der Voraussetzung, daß überall $K(s, t) > 0$ gilt, läßt sich leicht zeigen, daß λ_1 einfacher Eigenwert sein muß. Sonst gäbe es dazu nämlich mindestens eine das Zeichen wechselnde normierte Eigenfunktion $\varphi_1(s)$. Man bilde dann mit kleinem $\varepsilon > 0$ den positiven Kern

$$(58) \quad K_0(s, t) = K(s, t) + \varepsilon \varphi_1(s) \cdot \varphi_1(t).$$

Dieser hat die gleichen Eigenfunktionen wie $K(s, t)$ zu den gleichen Eigenwerten, nur gehört jetzt $\varphi_1(s)$ zum Eigenwert

$$\frac{\lambda_1}{1 + \lambda_1 \varepsilon} < \lambda_1.$$

Nach III müßte hierzu aber eine durchweg nicht negative Eigenfunktion gehören, was zum Widerspruch führt.

V. Setzt man nur (53) voraus, so soll jetzt gezeigt werden, daß zum ersten positiven Eigenwert wenigstens eine nicht negative Eigenfunktion $\varphi_1(s)$ gehört. Nach II braucht nur noch die Annahme berücksichtigt zu werden, daß λ_1 mindestens zweifacher Eigenwert ist. Es gibt also zu λ_1 auch eine Eigenfunktion $\varphi_2(s)$, für die

$$(59) \quad \int \varphi_2(s) ds = 0$$

ist. $\varphi_2(s)$ ist damit auch Eigenfunktion für alle Kerne

$$(60) \quad K_\varepsilon(s, t) = K(s, t) + \varepsilon > 0 \quad (\varepsilon > 0).$$

Gemäß IV gilt für die ersten positiven Eigenwerte $\lambda_{1\varepsilon}$ dieser Kerne

$$\lambda_{1\varepsilon} < \lambda_1.$$

Daher kann man aus den zugehörigen nichtnegativen normierten Eigenfunktionen $\varphi_{1\varepsilon}(s)$, die ja gleichgradig stetig und gleichmäßig beschränkt sind (Schwarzsche Ungleichung!), eine konvergente Teilfolge zu konvergenten $\lambda_{1\varepsilon}$ aussondern; ihre Grenzfunktion sei $\Phi(s) \geq 0$, der Grenzwert der $\lambda_{1\varepsilon}$ sei $\bar{\lambda}_1 \leq \lambda_1$.

Offenbar gilt

$$\Phi(s) - \bar{\lambda}_1 \int K(s, t) \Phi(t) dt = 0,$$

woraus $\bar{\lambda}_1 = \lambda_1$ folgt. Zum kleinsten positiven Eigenwert gehört somit mindestens eine nichtnegative Eigenfunktion. Wie unter III folgt, daß diese im Falle des Bestehens von (54) im Innern von \mathfrak{B} wirklich positiv ist.

VI. Sei im folgenden r ein beliebiger Randpunkt von \mathfrak{B} . Wir wollen zunächst annehmen

$$(61) \quad K(r, t) \neq 0$$

als Funktion von t für jeden Randpunkt r .

Setzt man in (57) $s_1 = r$, so ist sofort zu sehen, daß die nichtnegative Eigenfunktion $\varphi_1(s)$ am Rande nicht verschwinden kann. Damit ist also für alle s

$$(62) \quad \varphi_1(s) > 0.$$

Nun gibt es eine negative untere Grenze $-m$ des Quotienten

$$(63) \quad Q(s) = \frac{\varphi_2(s)}{\varphi_1(s)},$$

wenn unter $\varphi_2(s)$ eine zweite zu $\varphi_1(s)$ orthogonale Eigenfunktion zu λ_1 verstanden wird. Damit ist $\varphi(s) = m\varphi_1(s) + \varphi_2(s) \geq 0$, mithin eine nichtnegative Eigenfunktion, die auch den Wert Null annimmt; daß diese Nullstelle nicht im Innern von \mathfrak{B} liegt, folgt wieder aus (57), daß sie nicht auf dem Rande

liegen kann, sagen die zu (62) führenden Überlegungen aus. Wir erhalten so aus der Annahme der Existenz eines $\varphi_2(s)$ einen Widerspruch.

VII. Jetzt soll die Annahme (61) eliminiert werden. \mathfrak{R}_1 sei die Gesamtheit der Randpunkte, für die

$$(64) \quad K(r, t) \equiv 0; \quad r < \mathfrak{R}_1$$

gilt. Um jeden Punkt $r < \mathfrak{R}_1$ schlagen wir eine Kugel (von der Dimensionenzahl von \mathfrak{B}) vom Radius δ . \mathfrak{B}_1 enthalte alle Punkte von \mathfrak{B} , die in einer dieser Kugeln liegen. Der Inhalt von \mathfrak{B}_1 ist ersichtlich kleiner als $A\delta$, wo A eine von δ unabhängige Zahl ist. — m bezeichne jetzt die untere Grenze von $Q(s)$ aus (62) im Restbereich $\mathfrak{B}_2 = \mathfrak{B} - \mathfrak{B}_1$. Dort ist also $m\varphi_1 + \varphi_2 \geq 0$, wobei das Gleichheitszeichen für mindestens einen Punkt von \mathfrak{B}_2 (eventuell auf dem Rand) gilt. In \mathfrak{B}_1 ist $\lambda_1 K(s, t) < \varepsilon$, wo s mit δ nach 0 strebt. Wegen

$$\varphi_2(s) = \lambda_1 \int K(s, t) \varphi_2(t) dt$$

erhält man daher mittels der Schwarzschen Ungleichung für Punkte $s < \mathfrak{B}_1$

$$(65) \quad |\varphi_2(s)| \leq \varepsilon \int |\varphi_2(t)| dt \leq \varepsilon \sqrt{B} \sqrt{\int \varphi_2^2(t) dt} = \varepsilon \sqrt{B},$$

wenn B den Inhalt von \mathfrak{B} bedeutet. Mithin ist dort

$$(66) \quad m\varphi_1(s) + \varphi_2(s) \geq -\varepsilon \sqrt{B}.$$

Nun betrachten wir

$$\begin{aligned} m\varphi_1(s) + \varphi_2(s) &= \lambda_1 \int_{\mathfrak{B}_1} K(s, t) [m\varphi_1(t) + \varphi_2(t)] dt \\ &\quad + \lambda_1 \int_{\mathfrak{B}_2} K(s, t) [m\varphi_1(t) + \varphi_2(t)] dt. \end{aligned}$$

Da, der zweite Summand rechter Hand positiv ist, gilt unter Beachtung von (66) in \mathfrak{B}_1

$$m\varphi_1(s) + \varphi_2(s) \geq -\sqrt{B}\varepsilon(A\delta).$$

Fortgesetzte Wiederholung dieser Abschätzung liefert in \mathfrak{B}_1

$$m\varphi_1(s) + \varphi_2(s) \geq -\sqrt{B}\varepsilon(A\delta)^n,$$

also schließlich mit wachsendem n

$$m\varphi_1(s) + \varphi_2(s) \geq 0 \quad \text{in ganz } \mathfrak{B}.$$

Da dabei das Gleichheitszeichen für einen nicht zu \mathfrak{R}_1 gehörenden Punkt zutrifft, erhalten wir den gleichen Widerspruch wie unter VI. Damit ist also ganz allgemein das Vorhandensein von mehr als einer Eigenfunktion zum kleinsten positiven Eigenwert λ_1 als widerspruchsvoll nachgewiesen.

VIII. Jetzt soll gezeigt werden, daß es überhaupt einen positiven Eigenwert λ_1 gibt. Zu diesem Zweck betrachten wir den iterierten Kern

$$(67) \quad K_2(s, t) = \int K(s, \tau) K(\tau, t) dt,$$

der die gleichen Eigenschaften (53) und (54) besitzt wie $K(s, t)$. Zu ihm gehören die gleichen Eigenfunktionen wie zu $K(s, t)$, die Eigenwerte sind die Quadrate der ursprünglichen Eigenwerte. Da jeder symmetrische Kern mindestens einen Eigenwert besitzt, gibt es einen kleinsten Eigenwert λ_1^2 zu $K_2(s, t)$, der nach VII einfach zählt und zu dem eine positive Eigenfunktion $\varphi_1(s)$ gehört. Dann ist $\varphi_1(s)$ Eigenfunktion von $K(s, t)$ zum Eigenwert $\lambda_1 > 0$, da für $-\lambda_1 < 0$ die eventuell zugehörige Eigenfunktion nicht durchweg positiv sein könnte. Gleichzeitig zeigt dieser Schluß, daß $-\lambda_1$ überhaupt nicht Eigenwert von $K(s, t)$ sein kann.

Damit ist unser Satz in vollem Umfang bewiesen.

Bemerkung. Da sich bei weiterer Iteration des Kernes stets die Eigenschaften (53) und (54) übertragen, läßt sich der Satz auch aufrechterhalten, falls $K(s, t)$ selbst nicht stetig ist, wohl aber ein n -ter iterierter Kern $K_n(s, t)$.

§ 7.

Der Einschließungssatz bei Kernen mit positiven Eigenfunktionen.

In engem Anschluß an C. 11 zeigen wir noch folgenden einfachen

Satz. Ist die zur Bildung der Funktion $G(s)$ nach (8) verwendete Funktion $F_1(s)$ [vgl. (2)] positiv und besitzt $K(s, t)$ eine positive Eigenfunktion $\varphi_r(s)$ zum Eigenwert λ_r , so liegt λ_r zwischen dem Maximum und dem Minimum von $G(s)$.

Beweis. Aus

$$F_0(s) - G(s) \int K(s, t) F_0(t) dt = 0$$

folgt

$$F_0(s) - \lambda_r \int K(s, t) F_0(t) dt = (G(s) - \lambda_r) \int K(s, t) F_0(t) dt$$

und daher

$$(68) \quad \int \varphi_r(s) (G(s) - \lambda_r) F_1(s) ds = 0.$$

Da alle übrigen Faktoren positiv sind, muß hiernach $G(s) - \lambda_r$ das Zeichen wechseln.

Nimmt man das Ergebnis von § 6 hinzu, so folgt der

Zusatz: Ist neben $F_1(s) \geq 0$ (53) und (54) erfüllt, so liegt zwischen Größtwert und Kleinstwert von $G(s)$ der absolut kleinste Eigenwert λ_1 ; dieser ist positiv und einfach, $-\lambda_1$ ist kein Eigenwert.

(Eingegangen am 1. 7. 1941.)

Inversion und konforme Abbildung von Komplementärgebieten.

Von

Ernst Richard Neumann in Marburg.

In zwei früheren Arbeiten habe ich das folgende Problem behandelt: Angenommen, man kenne die Riemannsche Abbildungsfunktion (Abbildung auf den Einheitskreis) für eines der beiden von einer Jordankurve σ begrenzten Gebiete — gesucht wird die Abbildungsfunktion für das andere.

Unter gewissen über die Kurve gemachten einschränkenden Voraussetzungen (nämlich stetiger Biegung und endlicher Krümmung) habe ich dieses Problem in jenen Arbeiten gelöst¹⁾. Ich führte nämlich die potential-theoretischen Randwertprobleme und damit auch die Abbildungsaufgaben für das Innen- und für das Außengebiet einer geschlossenen Kurve zurück auf die Ermittlung der *gleichen* zwei Belegungen dieser Kurve σ selbst, nämlich der „natürlichen Belegung“ und der „Grundrestbelegung“, und dann zeigte ich weiter, wie man diese beiden Belegungen bestimmen kann, wenn man nur die Abbildungsfunktion eines der beiden Gebiete kennt. — Ich wies ferner schon darauf hin, daß man sogar die zweite, die Grundrestbelegung, entbehren kann, wenn man speziell den Übergang vom Innen- zum Außengebiet vollziehen will [vgl. II, S. 699].

In dieser Mitteilung will ich nun zeigen, daß man bei diesen ebenen Problemen²⁾ überhaupt die Grundrestbelegung gänzlich ausschalten kann, wenn man noch das Hilfsmittel der Inversion zu Hilfe nimmt — gleichgültig, ob man vom Innen- zum Außengebiet übergehen will oder umgekehrt.

Danach konzentriert sich dann das Interesse im wesentlichen auf die „natürliche Belegung“ der Kurve σ , d. h. die Gleichgewichtsverteilung eines elektroiden Fluidums (von der Masse 1) ohne Einwirkung äußerer Kräfte, und das einzige, was ich aus den früheren Arbeiten benutze, und was ich daher hier kurz rekapitulieren möchte, ist das Verfahren, wie man die Dichtigkeit dieser natürlichen Belegung einer Kurve σ ermittelt, wenn eine Abbildungsfunktion ihres Innengebietes \mathfrak{J} gegeben ist:

¹⁾ Math. Annalen 116, S. 534—554 und S. 664—700 — im folgenden kurz mit I und II zitiert. — An den erwähnten Voraussetzungen über die Kurve σ sei auch im folgenden immer festgehalten.

²⁾ Im Raume liegen die Verhältnisse wesentlich anders: Da sehe ich bisher keinen Weg, der ohne Benutzung der Grundrestbelegung einen Zusammenhang zwischen den Randwertproblemen von Innen- und Außenraum herstellt.

Diese als bekannt vorausgesetzte Abbildungsfunktion sei

$$(1) \quad z = \varphi(w),$$

— genauer eine Funktion, welche den Einheitskreis der w -Ebene auf das Gebiet \mathfrak{J} abbildet. Dann kann man zunächst mittels dieser Funktion ein System von gewissen reellen Größen J_{nx} einführen, oder eine Matrix

$$(2) \quad (\mathbf{J}) = \begin{pmatrix} J_{01} & J_{02} & J_{03} & \dots \\ J_{11} & J_{12} & J_{13} & \dots \\ J_{21} & J_{22} & \dots & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix},$$

und zwar sind diese Größen J_{nx} im wesentlichen die Real- und Imaginärteile der Koeffizienten in der Entwicklung des Logarithmus des Differenzenquotienten von $\varphi(w)$, also von

$$(3i) \quad \log \frac{\varphi(w_1) - \varphi(w_2)}{w_1 - w_2},$$

in eine Doppelpotenzreihe der Argumente w_1 und w_2 (wegen der Einzelheiten sei auf II, S. 674 verwiesen). — Ferner vermittelt jene Funktion (1) nach bekannten Prinzipien [vgl. z. B. II, S. 666] die Kenntnis derjenigen Greenschen Funktion $\mathfrak{G}_{(o)}$ des Gebietes \mathfrak{J} , deren Pol $o = \varphi(0)$ ist. Die in Punkten s auf σ anzutreffenden Werte des zu dieser Greenschen Funktion $\mathfrak{G}_{(o)p} = \lambda(x, y)$ konjugierten Potentials $\omega(x, y)$ liefern dann auf σ einen Parameter, der sich bei einer Umlaufung von σ monoton um 2π ändert.

Als Funktion dieses Parameters soll dann die Dichtigkeit ϱ_s der natürlichen Belegung von σ bestimmt, nämlich durch eine Fouriersche Reihe dargestellt werden. Freilich wollen wir nicht die Fouriersche Reihe für ϱ selber suchen, was an sich naheläge, sondern für den Quotienten zwischen ϱ und einer der Abbildungsfunktion (1), und damit den zugehörigen „Greenschen Koordinaten λ und ω “ [vgl. II, S. 669] charakteristisch zugehörigen Funktion, nämlich der Dichtigkeit

$$(4) \quad \eta_{(o)s} = \frac{1}{2\pi} \frac{\partial \lambda}{\partial v_s} \quad [\text{vgl. II, S. 671}]$$

der dem Koordinatenpole o entsprechenden „Greenschen Belegung“ [vgl. I, S. 541]³⁾. — Wir machen demnach für die Dichtigkeit ϱ_s der natürlichen Belegung den Ansatz

$$(5) \quad \varrho_s = \eta_{(o)s} \left(r_0 + \sum_1^{\infty} (v) r_p(\omega) \right), \quad (r_0 = \int \varrho_\sigma d\sigma = 1),$$

³⁾ Man könnte diese Belegung zweckmäßig etwa als „Koordinatenbelegung“ bezeichnen. — In allen Punkten s von σ ist ihre Dichtigkeit $\eta_{(o)s} > 0$.

wo die „Kreisfunktionen $p_v(x)$ “ die in geeigneter Weise normierten Funktionen $\sin(v)x$ und $\cos(v)x$ sind und das Symbol (v) in der Bedeutung steht:

$$(5') \quad (v) = \left[\frac{v+1}{2} \right] \quad [\text{vgl. II, S. 674}].$$

Wir bezeichnen (5) als die „Robin-Entwicklung“ der Funktion ϱ und die Koeffizienten r_n als ihre „Robin-Koeffizienten“. — Diese allein gilt es dann also noch zu bestimmen — und hierfür habe ich die folgende Regel abgeleitet: Man führe noch die Matrix (J) der „nullfreien“ Größen J_{rs} ein, d. h. die symmetrische Matrix der von (\hat{J}) bei Streichung der ersten Zeile verbleibenden Elemente:

$$(6) \quad (J) = \begin{pmatrix} J_{11} & J_{12} & J_{13} & \dots \\ J_{21} & J_{22} & J_{23} & \dots \\ J_{31} & J_{32} & \dots & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}$$

und bilde weiter die Matrizen

$$(7) \quad (\hat{J}) \{1 + (J) + (J)^2 + \dots + (J)^\mu\} \quad \text{für } \mu = 1, 2, 3 \dots$$

Als dann sind die gesuchten Robin-Koeffizienten r_μ in (5) die Grenzwerte, denen sich bei fortgesetzter Vergrößerung von μ die Elemente der ersten Zeile dieser Matrizen nähern. — Damit können wir die natürliche Belegung ϱ als aus der Abbildungsfunktion des Innengebietes \mathfrak{J} bestimmt ansehen.

Nach dieser Rekapitulation wenden wir uns nun dem Hauptproblem zu: Wir nehmen jetzt also die Abbildungsfunktion $\varphi(w)$ für eines der beiden von unserer Kurve σ begrenzten Gebiete als bekannt an, gleichgültig ob dieses das Innen- oder das Außengebiet ist. Dieses ausgezeichnete Gebiet nennen wir \mathfrak{J} und sein Komplementärgebiet \mathfrak{R} (wie schon in II, S. 699). — Dem Gebiet \mathfrak{J} kann man mittels seiner Abbildungsfunktion $\varphi(w)$ ein System von charakteristischen Zahlen $Z_{n,x}$ zuordnen, und zwar verstehen wir unter ihnen, falls \mathfrak{J} das Innengebiet \mathfrak{J} ist, die schon oben eingeführten Größen $J_{n,x}$, wenn aber \mathfrak{J} das Außengebiet \mathfrak{R} darstellt, sind unter den $Z_{n,x}$ die Größen $A_{n,x}$ zu verstehen, die aus der Entwicklung von

$$(3a) \quad \log \frac{(w_2 - c) \varphi(w_1) - (w_1 - c) \varphi(w_2)}{w_1 - w_2} \quad [\text{vgl. II, S. 678}]$$

ebenso hervorgehen wie oben die $J_{n,x}$ aus der Entwicklung der Funktion (3i). Dabei sind $\varphi(w)$ und c erklärt durch die Beziehung

$$(3'a) \quad \varphi(w) = \frac{\varphi(w)}{w - c}, \quad (|c| < 1)$$

welche die bekannte polare Unstetigkeit der Abbildungsfunktion $\varphi(w)$ des Außengebietes zum Ausdruck bringt.

Aus den so dem Gebiete 3 zugeordneten Größen Z_{nx} bilden wir nun wieder zwei Matrizen:

$$(8) \quad (\dot{Z}) = \begin{pmatrix} Z_{01} & Z_{02} & Z_{03} & \dots \\ Z_{11} & Z_{12} & Z_{13} & \dots \\ Z_{21} & Z_{22} & \dots & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \quad \text{und} \quad (Z) = \begin{pmatrix} Z_{11} & Z_{12} & Z_{13} & \dots \\ Z_{21} & Z_{22} & Z_{23} & \dots \\ Z_{31} & Z_{32} & \dots & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}.$$

Weiter nehmen wir in dem Komplementärgebiet \mathfrak{R} von 3 einen beliebigen Punkt k an und bilden die z -Ebene mit den beiden Gebieten 3 und \mathfrak{R} durch Inversion mit dem Zentrum k auf eine \bar{z} -Ebene ab. Dann ist das zu 3 konjugierte Gebiet $\bar{3}$ sicher das Innengebiet der zu σ konjugierten Kurve $\bar{\sigma}$, denn wenn eben k nicht zu 3 gehört, so gehört auch der zu k konjugierte Punkt, d. h. der Punkt ∞ , nicht zu $\bar{3}$.

Mit dem Übergang von σ zu $\bar{\sigma}$ gehen gleichzeitig aber auch alle Greenschen Niveaulinien und Radiallinien, d. h. alle Koordinatenlinien im Gebiete 3 [vgl. II, S. 668] über in entsprechende Kurven im Gebiete $\bar{3}$, und es haben sogar korrespondierende Punkte von 3 und $\bar{3}$ gleiche Greensche Koordinaten, wenn wir nur zu Polen o und o' der Systeme konjugierte Punkte wählen, und ganz entsprechendes gilt natürlich auch von den Gebieten \mathfrak{R} und $\bar{\mathfrak{R}}$. — Diese Überlegungen in Verbindung mit (4) führen leicht zu folgendem ganz allgemeinen Resultat: Zwischen den Dichtigkeiten der zu irgend zwei konjugierten Polen p und p' gehörigen Greenschen Belegungen von σ bzw. $\bar{\sigma}$ besteht in konjugierten Punkten s und s' die Beziehung

$$(9) \quad r_s \eta_{(p)s} = r_{s'} \bar{\eta}_{(p')s'},$$

also für die „Koordinatenbelegungen“:

$$(9') \quad r_s \eta_{(o)s} = r_{s'} \bar{\eta}_{(o')s'},$$

wenn r stets den Abstand vom Inversionszentrum k bedeutet⁴⁾.

Nach diesen mehr allgemeinen Vorbemerkungen wollen wir jetzt daran gehen, die Zahlen $\bar{Z}_{nx} \equiv J_{nx}$ des Gebietes $\bar{3}$ zu berechnen: Die Abbildungsfunktionen φ und $\bar{\varphi}$ der Gebiete 3 und $\bar{3}$ auf den Einheitskreis der w -Ebene hängen in folgender Weise zusammen:

$$(10) \quad \bar{\varphi}(w) - k = \frac{1}{\varphi(w) - k} \quad (\text{Ausdruck der Inversion}),$$

⁴⁾ Diese Beziehung hat schon C. Neumann, freilich auf einem wesentlich anderen Wege, bewiesen [vgl. Leipz. Ber. 62. (1910), Formel (A), S. 154], denn seine „induzierte Belegung $y_s^{(p)}$ “ ist eben nichts anderes wie unsere „Greensche Belegung $\eta_{(p)s}$ “ [vgl. I, S. 541, Anm. 7].

Daraus ergibt sich:

$$(11i) \quad \log \frac{\tilde{\varphi}(w_1) - \tilde{\varphi}(w_2)}{w_1 - w_2} = \log \frac{\varphi(w_1) - \varphi(w_2)}{w_1 - w_2} + \pi i \\ - \log(\varphi(w_1) - k) - \log(\varphi(w_2) - k),$$

und hieraus ersehen wir, daß von den Entwicklungskoeffizienten der Logarithmen der beiden Differenzenquotienten *alle nullfreien übereinstimmen*, denn die Zusatzfunktionen rechterhand hängen lediglich von je einer der Variablen w_1 und w_2 ab, liefern also bei der Doppelreihenentwicklung nur Beiträge zu den Gliedern mit Koeffizienten mit mindestens einem Index 0. — Diese Übereinstimmung überträgt sich von den Entwicklungskoeffizienten auf die Größen Z und \tilde{Z} [vgl. II, S. 674]. — Dabei ist allerdings zunächst vorausgesetzt, daß beide Gebiete 3 und $\tilde{3}$ Innengebiete sind, was ja bezüglich 3 nicht zuzutreffen braucht; trifft es zu, so ist also

$$(12i) \quad \tilde{Z}_{rx} = J_{rx}, \quad \text{sobald} \quad v, x \neq 0,$$

ist aber 3 das Außengebiet von σ , so erhalten wir unter Berücksichtigung von (3'a):

$$(11a) \quad \log \frac{\tilde{\varphi}(w_1) - \tilde{\varphi}(w_2)}{w_1 - w_2} = \log \frac{(w_2 - c) \psi(w_1) - (w_1 - c) \psi(w_2)}{w_1 - w_2} + \pi i \\ - \log(\psi(w_1) - k(w_1 - c)) - \log(\psi(w_2) - k(w_2 - c)),$$

woran sich ganz entsprechende Schlüsse knüpfen lassen wie oben an (11i). Das Resultat ist:

$$(12a) \quad \tilde{Z}_{rx} = A_{rx}, \quad \text{sobald} \quad v, x \neq 0.$$

Also stimmen stets die nullfreien Größen Z und \tilde{Z} überein:

$$(13) \quad \tilde{Z}_{rx} = Z_{rx}, \quad \text{sobald} \quad v, x \neq 0,$$

oder: Die ganze Matrix der nullfreien Größen Z verhält sich der Inversion gegenüber invariant:

$$(13^*) \quad (\tilde{Z}) = (Z).$$

Auch die Änderungen, welche die Elemente Z mit einem Index 0 erfahren, kann man den Gleichungen (11i) und (11a) sogleich entnehmen: Es ist allgemein

$$(14) \quad \tilde{Z}_{0x} = Z_{0x} - \mathfrak{z}_x(k), \quad x = 1, 2, 3 \dots$$

wo die $z_x(k)$ aus den Entwicklungskoeffizienten der Funktionen

$$(15i) \quad \log(\varphi(w) - k)$$

bzw.

$$(15a) \quad \log(\psi(w) - k(w - c))$$

ebenso entstehen, wie die J_{ox} und A_{ox} , aus denen der Funktionen (3i) bzw. (3a) hervorgingen [vgl. II, S. 674 (17)].

Nachdem wir so über das ganze System der dem Gebiete $\tilde{3}$ charakteristischen Größen $\tilde{Z}_{nx} \equiv \tilde{J}_{nx}$ Bescheid wissen, können wir in der Robin-Entwicklung der Dichtigkeit $\tilde{\varrho}$ der natürlichen Belegung der Kurve $\tilde{\sigma}$

$$(16) \quad \tilde{\varrho}_s = \tilde{\eta}_{(\omega')s} \left(1 + \sum_1^{\infty} (\nu) \tilde{r}_r p_r(\tilde{\omega}_s) \right)$$

die Koeffizienten \tilde{r}_r nach den obigen Regeln sofort angeben und damit $\tilde{\varrho}$ bestimmen. — Dieses $\tilde{\varrho}$ kann man nun aber auch ansehen als die Dichtigkeit der speziellen Pole ∞ entsprechenden Greenschen Belegung von $\tilde{\sigma}^5$. Gehen wir also jetzt von $\tilde{\sigma}$ zur ursprünglichen Kurve σ zurück, so liefert (9):

$$r_s \tilde{\varrho}_s = r_s \tilde{\eta}_{(\infty)s} = r_s \eta_{(k)s}$$

und (9'):

$$r_s \tilde{\eta}_{(\omega')s} = r_s \eta_{(\omega)s}$$

und daraus ergibt sich in Verbindungen mit (16), da ja $\tilde{\omega}_s = \omega_s$ ist, das Resultat:

$$\eta_{(k)s} = \eta_{(\omega)s} \left(1 + \sum_1^{\infty} (\nu) \tilde{r}_r p_r(\omega_s) \right).$$

So haben wir denn die Greensche Belegung der Kurve σ bestimmt, wie sie einem Punkte k des Komplementärgebietes R von 3 als Pol entspricht — vorausgesetzt, daß wir die Abbildungsfunktion des Gebietes 3 kennen.

Zusammenfassung.

Ist eine Abbildungsfunktion $\varphi(w)$ des Innen- oder Außengebietes einer Jordankurve σ bekannt, und ist o derjenige Punkt dieses so ausgezeichneten Gebietes 3 , dem der Anfangspunkt $w = 0$ der w -Ebene entspricht, so kann man nach früher auseinandergesetzten Prinzipien dem Gebiete 3 ein System von Zahlen Z_{nx} zuordnen (nämlich das System der Größen J_{nx} bzw. A_{nx} meiner Arbeit II, vgl. S. 674 und 678), und zugleich erhalten wir durch Abbildung der Kreise $|w| = \text{const}$ und der Radiivektoren $\arg w = \text{const}$ in 3 ein Greensches Koordinatensystem λ, ω (mit dem Pole o) und damit auch einen Parameter ω_s zur Bestimmung von Punkten s auf σ .

⁵⁾ Das folgt aus den Formeln (14) und (16) in I, S. 541 und 542, wird aber z. B. auch bestätigt durch die Formeln (30) bzw. (30*) daselbst S. 549 und 550.

Als dann gilt für die Dichtigkeit der zu einem beliebigen Punkte k des Komplementärgebietes \mathfrak{R} von \mathfrak{J} als Pol gehörigen Greenschen Belegung von σ als Funktion dieses Randparameters ω die Robin-Entwicklung

$$(17) \quad \eta_{(\omega),s} = \eta_{(\omega),s} \left(1 + \sum_{\nu=1}^{\infty} \bar{\tau}_{\nu}(k) p_{\nu}(\omega_s) \right),$$

wo die p_{ν} die Kreisfunktionen sind [vgl. II, S. 674] und die $\bar{\tau}_{\nu}(k)$ die Grenzwerte bedeuten, denen sich die Elemente der ersten Zeile in den Matrizen

$$(\tilde{Z}) (1 + (Z) + (Z)^2 + \dots + (Z)^{\mu}) \quad \mu = 1, 2, 3 \dots$$

mit wachsendem μ nähern. — Dabei bedeutet (Z) die Matrix der null/freien Größen $Z_{n,x}$ und die Matrix (\tilde{Z}) entsteht aus ihr, wenn man sie noch durch Darübersetzen einer weiteren Zeile ergänzt:

$$(\tilde{Z}) = \begin{pmatrix} Z_{01} - \beta_1(k) & Z_{02} - \beta_2(k) & Z_{03} - \beta_3(k) & \dots \\ Z_{11} & Z_{12} & Z_{13} & \dots \\ Z_{21} & Z_{22} & \dots & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix},$$

wo die β_{ν} in den bei (15i) bzw. (15a) angegebenen Bedeutungen stehen. (Sie sind die einzigen in der Formel (17) auftretenden Größen, die von der Lage von k abhängen.)

Mit der Greenschen Belegung (17) von σ , wie sie einem Pole k aus \mathfrak{R} entspricht, ist aber auch die *Greensche Funktion* dieses Komplementärgebietes \mathfrak{R} mit dem Pole k bekannt [vgl. I, S. 541] und damit auch diejenige Abbildung von \mathfrak{R} auf den Einheitskreis der w -Ebene, bei der k in $w = 0$ übergeht⁶⁾. Diese Abbildung wird vermittelt durch die Funktion

$$w(z) = e^{i\alpha} \cdot e^{\Pi_k - \Gamma} (z - k) e^{\frac{-\int \eta(k) \log(z - \zeta) |d\zeta|}{v}} \quad [\text{vgl. II, S. 699}]$$

mit beliebig reellem α (Drehung). — Damit ist unser Problem gelöst.

Bemerkt sei noch, daß, wenn \mathfrak{R} das Innengebiet von σ ist, der Faktor $e^{\Pi_k - \Gamma}$ fortfällt (gleich 1 wird) — andernfalls sind Π_k und $\Gamma = \Pi_{\infty}$ als Potentiale der natürlichen Belegung zu berechnen, deren Dichtigkeit ϱ_s sich nach (5) und (7) ergibt.

⁶⁾ Wir können also bei diesem Verfahren, auch wenn wir von der Abbildung des Innen- zu der des Außengebietes übergehen, es von vornherein so einrichten, daß ein beliebiger äußerer Punkt $a (= k)$ in den Mittelpunkt des Einheitskreises der w -Ebene abgebildet wird, während das oben erwähnte in II, S. 700 angegebene Verfahren ohne Benutzung der Grundrestbelegung uns zunächst nur die „natürliche Abbildung“ des Außengebietes lieferte, bei der speziell der Punkt ∞ in $w = 0$ übergeht.

Anhang.

Über die Grundrestbelegung konjugierter Kurven in der Ebene
und konjugierter Flächen im Raume.

Oben machten wir in (9) Gebrauch davon, daß bei konjugierter Pol-Lage die Greenschen Belegungen konjugierter Kurven σ und $\bar{\sigma}$ (d. h. von Kurven, die durch Inversion oder Abbildung nach reziproken Radien auseinander hervorgehen) selbst „konjugierte Belegungen“ sind, wenn wir hierunter solche Belegungen verstehen, deren Dichtigkeiten c und \bar{c} in konjugierten Punkten s und s' in der Beziehung stehen:

$$(1E) \quad r_s c_s = r_{s'} \bar{c}_{s'} \quad (r = \text{Abstand vom Abbildungszentrum}).$$

Ähnlich kann man nach dem Zusammenhang zwischen den Grundrestbelegungen konjugierter Kurven fragen. — Die Antwort ist bereits in meinen Beiträgen gegeben (S. 109) und lautet, wenn man die in den Arbeiten I und II benutzte Definition der Grundrestbelegung zugrunde legt⁷⁾:

$$(2E) \quad r_s r_s \mathfrak{C}_{(s)s} = r_{s'} r_{s'} \bar{\mathfrak{C}}_{(s')s'},$$

wenn ebenso wie die Aufpunkte s und s' auch die Polpunkte s und s' konjugiert sind. Wir sagen dafür in wohl leicht verständlicher Weise: *Die Grundrestbelegungen konjugierter Kurven sind zweifach konjugiert.*

Wesentlich anders liegen die Verhältnisse bei konjugierten Flächen im Raume und dem Newtonschen Potential. Hier bezeichnet man als „konjugierte Belegungen“ solche, deren Dichtigkeiten c und \bar{c} in konjugierten Punkten s und s' in der Beziehung stehen:

$$(1R) \quad r_s^{\frac{3}{2}} c_s = r_{s'}^{\frac{3}{2}} \bar{c}_{s'} \quad [\text{vgl. z. B. Beiträge S. 110}]$$

und, wenn sie noch von je einem zweiten Punkte s bzw. s' abhängen, wären sie demnach als „zweifach konjugiert“ zu bezeichnen, wenn die Beziehung besteht:

$$(2R) \quad r_s^{\frac{3}{2}} r_s^{\frac{3}{2}} \mathfrak{C}_{(s)s} = r_{s'}^{\frac{3}{2}} r_{s'}^{\frac{3}{2}} \bar{\mathfrak{C}}_{(s')s'} \quad (s, s' \text{ und } s, s' \text{ konjugierte Punktepaare}).$$

Dann gilt aber nach Beiträge S. 116 nicht etwa, daß die Grundrestbelegungen konjugierter Flächen zweifach konjugiert zueinander sind, sondern: zweifach konjugiert ist (in der alten Bezeichnung, vgl. Anm. 7) die Grundrestbelegung der einen Fläche zu einer Belegung der anderen, deren Dichtigkeit sich von der Grundrestbelegung um ein angebbares Zusatzglied $-\frac{1}{4\pi^2} \Omega_{(s)s}$ unterscheidet: In unsere neue Bezeichnung umgesetzt, lautet das Resultat:

$$(3) \quad r_s^{\frac{3}{2}} r_s^{\frac{3}{2}} \left(\mathfrak{C}_{(s)s} + \frac{\partial_s c_s}{r} - \frac{1}{4\pi^2} \Omega_{(s)s} \right) = r_{s'}^{\frac{3}{2}} r_{s'}^{\frac{3}{2}} \left(\bar{\mathfrak{C}}_{(s')s'} + \frac{\bar{\partial}_{s'} \bar{c}_{s'}}{r} \right)$$

⁷⁾ Nach ihr ist die Dichtigkeit $\mathfrak{C}_{(s)s}$ der Grundrestbelegung das, was ich noch in meinen „Beiträgen“ [vgl. I, Anm. 1] mit $\mathfrak{C}_{(s)s} - \frac{\partial_s c_s}{r}$ bezeichnete [vgl. I, S. 548].

oder in symmetrischer Form geschrieben:

$$(3') \quad r_s^{\frac{3}{2}} r_s^{\frac{3}{2}} \left(\mathfrak{E}_{(s)s} + \frac{q_s q_s}{F} - \frac{1}{8\pi^2} \Omega_{(s)s} \right) = r_s^{\frac{3}{2}} r_s^{\frac{3}{2}} \left(\tilde{\mathfrak{E}}_{(s')s'} + \frac{\tilde{q}_{s'} \tilde{q}_{s'}}{F} - \frac{1}{8\pi^2} \tilde{\Omega}_{(s')s'} \right)$$

und für diese Funktion $\Omega_{(s)s}$ habe ich dort die Darstellung angegeben:

$$(4) \quad \Omega_{(s)s} = r_s \frac{\partial}{\partial v_s} \frac{\frac{1}{r_s} \frac{\partial}{\partial v_s} \frac{1}{E_{ss}}}{\frac{\partial v_s}{\partial v_s}} + r_s \frac{\partial}{\partial v_s} \frac{\frac{1}{r_s} \frac{\partial}{\partial v_s} \frac{1}{E_{ss}}}{\frac{\partial v_s}{\partial v_s}} - \frac{r_s r_s}{E_{ss}} \frac{\partial}{\partial v_s} \frac{\frac{1}{r_s} \frac{\partial}{\partial v_s} \frac{1}{E_{ss}}}{\frac{\partial v_s}{\partial v_s}}.$$

Danach ist diese Zusatzgröße Ω noch von der Lage des Abbildungszentrums abhängig, ein Umstand, der dem Resultat (3) oder (3') jede einfache Bedeutung nimmt. — Es soll nun hier gezeigt werden, daß sich dieser Übelstand beseitigen läßt.

Zunächst gestattet $\Omega_{(s)s}$ folgende Umformung:

$$(4') \quad \Omega_{(s)s} = \frac{1}{E_{ss}} \frac{\partial \log \frac{1}{E_{ss}}}{\partial v_s} \frac{\partial \log \frac{1}{E_{ss}}}{\partial v_s} - \frac{1}{E_{ss}} \frac{\partial \log \frac{r_s}{E_{ss}}}{\partial v_s} \frac{\partial \log \frac{r_s}{E_{ss}}}{\partial v_s},$$

und mittels der sich aus ähnlichen Dreiecken leicht ergebenden Beziehungen

$$\frac{r_s}{E_{ss}} = \frac{r_{s'}}{E_{s's'}} \quad \text{und} \quad \frac{r_s}{E_{ss}} = \frac{r_{s'}}{E_{s's'}}$$

folgt weiter mühelos für das zweite Glied:

$$(5) \quad r_s^{\frac{3}{2}} r_s^{\frac{3}{2}} \left(\frac{1}{E_{ss}} \frac{\partial \log \frac{r_s}{E_{ss}}}{\partial v_s} \frac{\partial \log \frac{r_s}{E_{ss}}}{\partial v_s} \right) = r_s^{\frac{3}{2}} r_s^{\frac{3}{2}} \left(\frac{1}{E_{s's'}} \frac{\partial \log \frac{1}{E_{s's'}}}{\partial v_{s'}} \frac{\partial \log \frac{1}{E_{s's'}}}{\partial v_{s'}} \right).$$

Unter Benutzung dieser Resultate (4') und (5) nimmt dann die Beziehung (3) die folgende Form an:

$$(6) \quad r_s^{\frac{3}{2}} r_s^{\frac{3}{2}} \left(\mathfrak{E}_{(s)s} + \frac{q_s q_s}{F} - \frac{1}{4\pi^2} \frac{1}{E_{ss}} \frac{\partial \log \frac{1}{E_{ss}}}{\partial v_s} \frac{\partial \log \frac{1}{E_{ss}}}{\partial v_s} \right) \\ = r_s^{\frac{3}{2}} r_s^{\frac{3}{2}} \left(\tilde{\mathfrak{E}}_{(s')s'} + \frac{\tilde{q}_{s'} \tilde{q}_{s'}}{F} - \frac{1}{4\pi^2} \frac{1}{E_{s's'}} \frac{\partial \log \frac{1}{E_{s's'}}}{\partial v_{s'}} \frac{\partial \log \frac{1}{E_{s's'}}}{\partial v_{s'}} \right),$$

und hier sind jetzt die zu \mathfrak{E} noch hinzutretenden Glieder tatsächlich unabhängig von der Lage des Abbildungszentrums.

Um das damit erhaltene Resultat einfach formulieren zu können, führen wir noch eine „ \mathfrak{F} -Belegung“ der Fläche σ ein, deren Dichtigkeit die nur von der Lage der Punkte s und s (Pol und Aufpunkt) abhängige Funktion

$$(7) \quad \mathfrak{F}_{(s)s} = \mathfrak{E}_{(s)s} + \frac{q_s q_s}{F} - \frac{E_{ss}}{4\pi^2} \frac{\partial}{\partial v_s} \frac{1}{E_{ss}} \frac{\partial}{\partial v_s} \frac{1}{E_{ss}}$$

ist, oder auch, was dasselbe:

$$(7') \quad \mathfrak{F}_{(s)s} = \mathfrak{E}_{(s)s} + \frac{\varrho_s \varrho_s}{r} - \frac{1}{4\pi^2} \cdot \frac{1}{E_{ss}} \cdot \frac{\cos(E_{ss}, \nu_s)}{E_{ss}} \cdot \frac{\cos(E_{ss}, \nu_s)}{E_{ss}}.$$

Dann besagt (6): Die \mathfrak{F} -Belegungen konjugierter Flächen σ und $\bar{\sigma}$ sind zu einander zweifach konjugiert:

$$(8) \quad r_s^{\frac{2}{3}} r_s^{\frac{2}{3}} \mathfrak{F}_{(s)s} = r_{s'}^{\frac{2}{3}} r_{s'}^{\frac{2}{3}} \mathfrak{F}_{(s')s'},$$

wo natürlich s, s' und s, s' Paare konjugierter Punkte (auf σ und $\bar{\sigma}$) sind.

Speziell bei der Kugelfläche ist $\mathfrak{F} \equiv 0$, es bestätigt sich also das Gesetz (8) in denkbar einfachster Weise. — Auch bei beliebigen Flächen σ dürfte \mathfrak{F} im Gegensatz zu \mathfrak{E} [vgl. Beitr. S. 82] immer stetig sein (auch bei Annäherung von s an den Pol s), und, da nach (7) mit der \mathfrak{F} -Belegung auch die \mathfrak{E} -Belegung (Grundrestbelegung) bekannt ist, dürfte sich danach diese \mathfrak{F} -Belegung als zweckmäßiger Ersatz für die Grundrestbelegung empfehlen.

Doch ist es mir bisher nicht gelungen, im allgemeinen Falle die Stetigkeit der Funktion \mathfrak{F} zu beweisen, und ebensowenig auch, eine selbständige Charakterisierung von \mathfrak{F} , etwa durch eine einfache Integralgleichung, zu finden.

(Eingegangen am 12. 9. 1940.)

Zur pythagoreischen Algebra: Quadratwurzel und Kubikwurzel.

Von

B. L. van der Waerden in Leipzig.

In der axiomatischen Begründung der pythagoreischen Musiklehre, über die Ptolemäus¹⁾ kurz referiert und die in Euklids *Sectio Canonis*²⁾ ausführlich dargestellt wird, spielt die Frage nach der Möglichkeit der Teilung eines Intervalles in zwei oder mehrere gleiche Teilintervalle eine große Rolle. Arithmetisch kommt sie auf die Frage hinaus, ob sich zwischen zwei Zahlen, die in einem gegebenen Verhältnis $a : b$ stehen, eine oder mehrere mittlere Proportionale einschieben lassen³⁾. Archytas von Taras, der hervorragendste unter den pythagoreischen Musiktheoretikern, hat bewiesen⁴⁾, daß eine solche Einschiebung unmöglich ist, wenn das Verhältnis $a : b$ ein „überteiliges“, d. h. wenn $a : b = n : (n + 1)$ ist.

Der Beweis des Archytas beruht, ebenso wie der in der *Sectio Canonis* dargestellte, auf folgendem Hilfssatz: Sind a' und b' die kleinsten Zahlen im Verhältnis $a : b$, so lassen sich ebenso viele mittlere Proportionale wie zwischen a und b auch zwischen a' und b' einschieben. Dieser ist ein Spezialfall des Satzes VIII 8 der *Elemente* von Euklid. Im gleichen VIII. Buch findet sich auch die vollständige Antwort auf die gestellte Frage: Die Einschiebung einer mittleren Proportionalen ist nur möglich, wenn a' und b' Quadratzahlen, von zwei mittleren Proportionalen, wenn sie Kubikzahlen sind (VIII 9 in Verbindung mit IX 8). Die Umkehrung (VIII 11 und 12) wird von Plato dem Pythagoreer Timaios in den Mund gelegt⁵⁾, wodurch bestätigt wird, daß es sich hier um pythagoreische Arithmetik handelt⁶⁾.

Wie sind nun die Pythagoreer zu der Frage nach den mittleren Proportionalen gekommen?

Ich glaube nicht, daß diese Frage aus der Musiktheorie erwachsen ist. Denn für diese sind zwar das arithmetische und das harmonische Mittel direkt wichtig (nämlich für die Teilung der Oktave in Quinte und Quarte), das geometrische aber nur indirekt, insofern die Nichthalbierbarkeit der über-

¹⁾ Klaudios Ptolemaios, *Harmonielehre* (ed. Düring), Buch I, Kap. 5.

²⁾ *Euclidis Opera Omnia* (ed. Heiberg u. Menge, Leipzig 1916), Bd. VIII.

³⁾ „Zahl“ bedeutet hier nach griechischem Sprachgebrauch „ganze Zahl“; die gesuchten mittleren Proportionalen sollen ebenfalls ganze Zahlen sein.

⁴⁾ Archytas bei Boetius, *De institutione musica* (ed. Friedlein, Leipzig 1867), S. 285–286.

⁵⁾ Timaios 31 B 4 und 32 A 7.

⁶⁾ Vgl. dazu P. Tannery, *Mém. scient.* III, Nr. 83, und K. Reidemeister, *Die Arithmetik der Griechen* (Leipzig 1939).

teiligen Intervalle im logischen Aufbau der Sectio Canonis eine Rolle spielt. Dieser Aufbau stellt aber eine Endstufe der Theorie dar, die erst auf Grund einer ausgebildeten Arithmetik möglich war.

Weit eher könnte das geometrische Mittel aus der Geometrie stammen. Zum ersten Male kommt die Konstruktion des geometrischen Mittels bei Euklid II 14 vor, wo es sich darum handelt, zu einem gegebenen Polygon ein flächengleiches Quadrat zu errichten. Diese Konstruktion ist ein unentbehrliches Hilfsmittel bei der Flächenanpassung (Euklid VI 28 und 29), die nach Eudem⁷⁾ eine Erfindung der Pythagoreer ist. Die Konstruktion von zwei geometrischen Mitteln zwischen zwei vorgegebenen Strecken ist äquivalent dem Problem, einen Würfel von gegebenem Rauminhalt zu konstruieren, und diese Äquivalenz bildet die Grundlage der stereometrischen Lösung des Problems der Würfelverdopplung von Archytas⁸⁾. Die geometrischen Mittel spielen also in der pythagoreischen Geometrie eine wichtige Rolle.

Trotzdem glaube ich nicht, daß die Wurzel der arithmetischen Problemstellung in der Geometrie zu suchen ist. Denn für die Pythagoreer sind die Zahlen immer das primäre; sie trieben Arithmetik, längst bevor es eine Geometrie gab, die sich an so schwere Probleme wie die Würfelverdopplung heranwagen konnte. Von der Geometrie aus gibt es auch keine plausible Erklärung, warum so ausgefallene Probleme wie die Flächenanpassung und die Würfelverdopplung eine so zentrale Rolle spielen. *Vielmehr sieht es so aus, als ob diese Probleme nur geometrische Verkleidungen von algebraischen Fragestellungen sind.* Die einfachsten nichtlinearen Gleichungen in einer oder zwei Unbekannten heißen nämlich

$$(1) \quad x^2 = a \quad \text{bzw.} \quad x^3 = a$$

bzw.

$$(2) \quad x \pm y = a, \quad xy = b,$$

und die geometrische Formulierung dieser Gleichungsaufgaben führt gerade auf die obenerwähnten Probleme.

Eine nähere Betrachtung der Arithmetik selbst führt zum gleichen Ergebnis. Bekanntlich hat es neben der offiziellen Arithmetik, die nur ganze Zahlen betrachtete, eine Logistik gegeben, die auch mit Brüchen rechnete. Die Arithmetik ersetzte diese Brüche durch Zahlenverhältnisse. So kann man die im VII. Buch Euklids entwickelte Theorie der auf ihre kleinsten Glieder reduzierten Zahlenverhältnisse als eine Theorie der gekürzten Brüche betrachten. Die Zusammensetzung von Verhältnissen entspricht der Multiplikation von Brüchen, die „Verdopplung“ eines Verhältnisses dem Quadrieren. Der Bildung der höheren Potenzen entspricht die arithmetische Theorie der geometrischen Reihen im VIII. und IX. Buch. So gesehen, ist

⁷⁾ Proklos in Euklid (ed. Friedlein), S. 419.

⁸⁾ Eudemos im Eutokioskommentar zu Archimedis Opera (Heiberg) III, S. 99.

die Frage nach den mittleren Proportionalen zwischen Zahlen nur die arithmetische Fassung der logistischen Frage nach der Quadratwurzel oder Kubikwurzel aus einer rationalen Zahl.

Nun wissen wir aber auch aus anderer Quelle, daß die Pythagoreer sich für die Berechnung von Quadratwurzeln interessierten. Theon v. Smyrna⁹⁾ und Proklos¹⁰⁾ geben uns das Bildungsgesetz der „Seiten- und Diagonalzahlen“, die der Pellschen Gleichung

$$2x^2 - y^2 = \pm 1$$

genügen und deren Verhältnisse Näherungen für $\sqrt{2}$ ergeben. Das Paar (5, 7) kommt bereits im Platonischen Staat (546 C) vor.

Zusammenfassend können wir die Genesis all dieser — auch in der *Epinomis*¹¹⁾ stark hervorgehobenen — Probleme so darstellen:

A. Algebraischer (logistischer) Ausgangspunkt:

1. Berechnung von Quadratwurzeln. 2. Berechnung von Kubikwurzeln.

B. Geometrische Einkleidung:

1. Quadratur des Rechtecks. 2. Kubatur eines Balkens (Würfelverdopplung).

C. Arithmetische Einkleidung:

1. Eine mittlere Proportionale. 2. Zwei mittlere Proportionale.

D. Musiktheoretisches Korollar:

1. Halbierung eines Intervalles.

Zur vollen Gewißheit, daß die hier betrachteten geometrischen und arithmetischen Probleme letzten Endes aus der Algebra stammen, gelangt man, wenn man auch noch die Keilschrifttexte heranzieht. Zahlreiche Quadratwurzel- und Kubikwurzeltabellen aus der Zeit von Isis I. bis Darius¹¹⁾ beweisen, wie wichtig diese Operationen den Babyloniern waren. Zahlreich sind auch die Probleme, die durch algebraische Umformung auf die Grundtypen (1) und (2) zurückgeführt wurden¹²⁾. Auch die Näherung $\frac{7}{5}$ für $\sqrt{2}$ findet man hier wieder. Daß die Pythagoreer bei ihren notorischen Beziehungen zum Orient das alles nicht gekannt, es aber in kaum mehr als einem Jahrhundert nebst „Pythagoreischem Lehrsatz“ und so vielem anderen keilschriftlich Überlieferten neu erfunden haben sollten, ist kaum glaublich.

⁹⁾ Theonis Smyrnaei Expos. rer. math. (ed. Hiller, Leipzig 1878), S. 43–44.

¹⁰⁾ Proklos, Komm. zum Staate Platos (ed. Kroll, 1901), S. 24–29.

¹¹⁾ Siehe O. Neugebauer, Mathematische Keilschrifttexte I (Quellen u. Studien z. Gesch. d. Math., Abt. A, Bd. 3, 1935), Kap. I, § 4.

¹²⁾ Siehe unter ¹¹⁾, für die Kubikwurzeln bes. S. 216 und 460–461.

¹³⁾ Siehe O. Toeplitz, Quellen und Studien B 2, S. 334.

Zu Hilberts 80. Geburtstag

Am 23. Januar 1942 ist David Hilbert 80 Jahre alt geworden. Die gesamte mathematische Welt grüßt ihn in Verehrung als den greisen Heros ihrer Wissenschaft, der durch die Tiefe seiner Gedanken und die Kraft seiner Persönlichkeit in dem vergangenen halben Jahrhundert als unbestrittener Führer der Forschung die Entwicklung der Mathematik bestimmt und gefördert hat. Die Großartigkeit seines Lebenswerkes und dessen Auswirkung in die Zukunft zu würdigen, ist nicht Aufgabe dieser Zeilen, wohl aber darf die Redaktion der Mathematischen Annalen, die Hilbert so lange Jahre geleitet hat, den Gefühlen herzlicher Dankbarkeit Ausdruck geben und dem Wunsche, daß ihm noch lange Jahre ruhevoller und glücklicher Muße beschieden sein mögen.

*Die Redaktion
der Mathematischen Annalen*

Über die Anzahl der Lösungen gewisser Aufgaben über Gitterpunktfiguren.

Von

Heinrich Tietze in München.

§ 1.

Einleitung.

1. Im Folgenden wird für drei und mehr Dimensionen eine Frage behandelt, deren Antwort im Falle von zwei Dimensionen bereits vorliegt (siehe Nr. 3, 4) und in der Lehre von den symmetrischen Funktionen eine bekannte Rolle spielt (Nr. 2).

Die Gitterpunkte, um die es sich bei unseren Fragen handelt, sind durchwegs solche mit *positiven* Koordinaten. Ist \mathfrak{M} eine endliche Menge von Gitterpunkten des n -dimensionalen Raumes mit den Koordinaten x_1, \dots, x_n , so möge mit $a_\mu^{(\nu)}$ die Anzahl der Gitterpunkte von \mathfrak{M} in der Ebene $x_\nu = \mu$ bezeichnet werden. Ist dann m der Grad von \mathfrak{M} , d. h. die Anzahl der Punkte von \mathfrak{M} , so erhalten wir für jedes $\nu = 1, 2, \dots, n$ ein System nicht-negativer ganzer Zahlen $\mathfrak{A}^{(\nu)} = \mathfrak{A}^{(\nu)}(\mathfrak{M}) = (a_1^{(\nu)}, a_2^{(\nu)}, \dots)$ mit der Summe m .

Geht man umgekehrt aus von n Systemen $\mathfrak{A}', \dots, \mathfrak{A}^{(n)}$ nichtnegativer ganzer Zahlen, alle mit der gleichen Summe m (≥ 1), so kann man fragen, ob es eine Gitterpunktmenge \mathfrak{M} vom Grad m gibt, so daß die zu \mathfrak{M} gehörigen Zahlensysteme $\mathfrak{A}'(\mathfrak{M}), \dots, \mathfrak{A}^{(n)}(\mathfrak{M})$ der Reihe nach mit den vorgegebenen Systemen $\mathfrak{A}', \dots, \mathfrak{A}^{(n)}$ übereinstimmen¹⁾. Weiter kann man nach der Anzahl $N(\mathfrak{A}' | \dots | \mathfrak{A}^{(n)})$ verschiedener derartiger Mengen \mathfrak{M} fragen²⁾. Für die Beantwortung dieser Fragen ist offenbar innerhalb jedes Zahlensystems $\mathfrak{A}^{(\nu)}$ die Reihenfolge der Zahlen $a_1^{(\nu)}, a_2^{(\nu)}, \dots$ belanglos, so daß man sie der Größe nach fallend geordnet annehmen darf: $a_1^{(\nu)} \geq a_2^{(\nu)} \geq \dots$. Auch sind natürlich nur

¹⁾ Vgl. Systeme von Partitionen und Gitterpunktfiguren I. Rekursionsformeln*). Sitz.ber. d. Bayer. Akad. d. Wiss., Jahrg. 1940 (S. 23–54), Nr. 3, S. 26. Wegen analoger Fragen im Kontinuierlichen vgl. ebenda, Jahrg. 1941, S. 19*, Sitzung vom 25. Okt. 1941, Punkt 5.

²⁾ Über diese Anzahlen und ihre Berechnung vgl. l. c. ¹⁾, S. 25ff. sowie „Gitterpunktfiguren“ III. Ein Satz über das Verhältnis der Lösungsanzahlen gewisser Partitionsaufgaben, IV. Formeln und Tabellen. Bayer. Sitz.ber., Jahrg. 1940, S. 133–145 und S. 147–166.

³⁾ Der Kürze halber soll weiterhin aus dieser Folge von Noten mit gleichem Ober-
titel die einzelne Note nur mit „Gitterpunktfiguren“, Nummer und Untertitel zitiert
werden oder überhaupt nur als „Note I“, „Note II“ usw.

die positiven unter den Zahlen $a_\mu^{(v)}$ wirklich von Bedeutung. Nennt man, wie üblich, ein System positiver ganzer Zahlen mit der Summe m , bei dem es auf die Reihenfolge der Zahlen nicht ankommt, eine Partition von m , so können wir nach dem Gesagten jedes Zahlensystem $\mathfrak{A}^{(v)}$ als eine Partition von m auffassen, und die Frage nach dem Vorhandensein von Gitterpunkt-mengen der gewünschten Art und nach ihrer Anzahl N hängt also wesentlich nur von den durch die Zahlensysteme $\mathfrak{A}^{(v)}$ dargestellten n Partitionen der Zahl m ab. Wir werden demnach künftighin von den „Partitionen $\mathfrak{A}^{(v)}$ “ sprechen und dabei die Zahlen $a_\mu^{(v)}$ mit $\mu > m$, die bei fallender Anordnung notwendig Null sind, weglassen; wir werden die Partition $\mathfrak{A}^{(v)}$ also in der Gestalt

$$\mathfrak{A}^{(v)} = (a_1^{(v)}, \dots, a_m^{(v)})$$

mit

$$(1) \quad a_1^{(v)} \geq a_2^{(v)} \geq \dots \geq a_m^{(v)} \geq 0, \quad a_1^{(v)} + \dots + a_m^{(v)} = m$$

annehmen.

Die folgenden Entwicklungen beschäftigen sich mit der Frage, zu welchen Systemen von Partitionen $\mathfrak{A}', \dots, \mathfrak{A}^{(n)}$ überhaupt wenigstens eine zugehörige n -dimensionale Gitterpunktmenge \mathfrak{M} vorhanden ist, also $N = N(\mathfrak{A}' | \dots | \mathfrak{A}^{(n)}) > 0$ ausfällt, wann ferner nur eine einzige Lösung \mathfrak{M} existiert, also $N = 1$ ist.

2. Nun ist im Falle $n = 2$ die Frage, wann für 2 Partitionen $\mathfrak{A}' = \mathfrak{A}$, $\mathfrak{A}'' = \mathfrak{B}$ die Anzahl $N = N(\mathfrak{A}' | \mathfrak{A}'') = N(\mathfrak{A} | \mathfrak{B}) > 0$ ausfällt, sowie im besonderen, wann $N = 1$ ist, vollständig beantwortet (vgl. Nr. 3). Dabei sind die für die Dimensionszahl $n = 2$ sich ergebenden Zahlen $N(\mathfrak{A} | \mathfrak{B})$ altbekannt wegen ihres Auftretens in der Lehre von den symmetrischen Funktionen. Sind

$$f_1 = \sum x_1, \quad f_2 = \sum x_1 x_2, \dots, \quad f_i = \sum x_1 x_2 \dots x_i$$

die symmetrischen Grundfunktionen der Variablen x_1, x_2, \dots, x_i , ist ferner $\mathfrak{A} = (a_1, a_2, \dots, a_m)$ eine Partition der natürlichen Zahl m und wird das Produkt $X_{\mathfrak{A}} = f_{a_1} f_{a_2} \dots f_{a_m}$ der Grundfunktionen (hierbei $f_0 = 1$ gesetzt) in der Gestalt

$$(2) \quad X_{\mathfrak{A}} = \sum_{\mathfrak{B}} C_{\mathfrak{A} | \mathfrak{B}} Y_{\mathfrak{B}}$$

dargestellt durch die verschiedenen Potenzproduktsommen

$$Y_{\mathfrak{B}} = \sum x_1^{b_1} x_2^{b_2} \dots x_m^{b_m},$$

wobei $\mathfrak{B} = (b_1, b_2, \dots, b_m)$ in der Summe (2) alle Partitionen von m durchläuft, dann ist jeder einzelne Koeffizient $C_{\mathfrak{A} | \mathfrak{B}}$ nichts anderes als die oben mit $N(\mathfrak{A} | \mathfrak{B})$ bezeichnete Zahl³⁾. Läßt man in (2) auch \mathfrak{A} alle Partitionen

³⁾ Vgl. Über symmetrische Funktionen von endlich oder abzählbar unendlich vielen Veränderlichen. Monatshefte f. Math. u. Phys. 48 (1939), S. 487–499; 49 (1940), S. 1–52; Nr. 1, S. 487, Nr. 10, Satz (a), S. 494, Nr. 17, S. 2.

von m durchlaufen, so erhält man in der Gesamtheit aller Koeffizienten $C_{\mathfrak{A}\mathfrak{B}} = N(\mathfrak{A}|\mathfrak{B})$ die bekannte Cayley-Perronsche Matrix.

Bei den diese Matrix verwendenden Beweisen des Fundamentalsatzes der Lehre von den symmetrischen Funktionen ist es wesentlich, daß man den Zeilen und Kolonnen der Matrix — anders gesagt den Partitionen \mathfrak{A} und den Partitionen \mathfrak{B} — eine solche Anordnung geben kann, daß in einer Diagonale lauter Einser, auf einer Seite der Diagonale lauter Nullen stehen. Für größere Werte von m kommen aber auch unter den Elementen auf der anderen Seite der Diagonale Nullen vor⁴⁾, jedoch keine Einser außerhalb der Diagonale⁵⁾. Wünscht man genau Bescheid darüber, an welchen Stellen der Cayley-Perronschen Matrix, d. h. im Koeffizientenschema der Gleichungen (2) der Wert 0, bzw. der Wert 1, bzw. Werte > 1 auftreten, so deckt sich, wie man sieht, diese Frage mit der zu Beginn dieser Nr. 2 genannten. Wir wollen sogleich besprechen, wie hier die Dinge liegen.

3. Die Antwort auf die eingangs der Nr. 2 über die Werte von $N(\mathfrak{A}|\dots|\mathfrak{A}^{(n)})$ im Falle $n = 2$ gestellte Frage — eine Antwort, die sich auf den Begriff der „konjugierten Partitionen“⁶⁾ und auf gewisse Ordnungsbeziehungen zwischen Partitionen⁷⁾ stützt — wird gegeben durch folgende Sätze⁸⁾:

(a) Zu den Partitionen $\mathfrak{A}, \mathfrak{A}'$ derselben Zahl m gibt es dann und nur dann wenigstens eine zweidimensionale Gitterpunktmenge \mathfrak{M} mit $\mathfrak{A}'(\mathfrak{M}) = \mathfrak{A}$, $\mathfrak{A}''(\mathfrak{M}) = \mathfrak{A}'$, d. h. es ist dann und nur dann $N(\mathfrak{A}'|\mathfrak{A}'') > 0$, wenn $\mathfrak{A}'' \leq (\mathfrak{A}')^{-1}$ — oder, was auf dasselbe hinausläuft, $\mathfrak{A}' \leq (\mathfrak{A})^{-1}$ — ist, wobei die zu einer Partition \mathfrak{A} konjugierte Partition mit \mathfrak{A}^{-1} bezeichnet wird.

⁴⁾ Nämlich stets, wenn $m \geq 6$ ist; vgl. l. c. ³⁾, S. 497, Anm. 23, sowie Nr. 24, Satz VII, S. 9.

⁵⁾ Vgl. den l. c. ⁴⁾ genannten Satz VII, sowie ebenda in Nr. 39–41 die Sätze (m), (n), (o), (p).

⁶⁾ Siehe etwa G. H. Hardy-E. M. Wright, An introduction to the theory of numbers, Oxford 1938, Chapter XIX, S. 272; vgl. auch l. c. ³⁾ Bd. 48, Nr. 11, S. 495 oder „Über Tripel konjugierter Partitionen“, Abhandlungen aus dem Math. Seminar der Hansischen Universität (Hamburg), Bd. 14 (1941), S. 273–284, Nr. 1.

⁷⁾ Von zwei verschiedenen Partitionen $\mathfrak{A} = (a_1, \dots, a_m)$, $\mathfrak{B} = (b_1, \dots, b_m)$ (jede monoton nicht-zunehmend angeordnet, vgl. (l), Nr. 1) wird \mathfrak{A} höher als \mathfrak{B} und \mathfrak{B} tiefer als \mathfrak{A} genannt ($\mathfrak{A} > \mathfrak{B}$, $\mathfrak{B} < \mathfrak{A}$), wenn $\sum_{i=1}^{\mu} a_i \geq \sum_{i=1}^{\mu} b_i$ für jeden Wert μ von 1 bis m

oder anders gesagt, wenn $\sum_{i=\mu}^m a_i \leq \sum_{i=\mu}^m b_i$ ($1 \leq \mu \leq m$) gilt. Vgl. l. c. ³⁾, Monatshefte, Bd. 48, Nr. 14, S. 497; Bd. 49, Nr. 24, S. 8, oder auch Note III, l. c. ³⁾, Nr. 5, S. 137, sowie wegen einer geometrischen Deutung Note VI. Konvexe Polygonzüge und Partitionen nebst deren Ordnungsbeziehungen, Bayer. Sitz.ber. 1941, S. 39–55.

⁸⁾ Siehe l. c. ³⁾, Bd. 49, Nr. 24, Satz VII, S. 9.

(b) Die Anzahl $N(\mathfrak{A}'|\mathfrak{A}'')$ der Gitterpunkt mengen \mathfrak{M} mit $\mathfrak{A}'(\mathfrak{M}) = \mathfrak{A}'$, $\mathfrak{A}''(\mathfrak{M}) = \mathfrak{A}''$ ist dann und nur dann $= 1$, wenn $\mathfrak{A}'' = (\mathfrak{A}')^{-1}$ [somit $\mathfrak{A}' = (\mathfrak{A}'')^{-1}$] ist. Die letztere Bedingung, daß \mathfrak{A}' und \mathfrak{A}'' zueinander konjugierte Partitionen sind, ist aber gleichbedeutend damit, daß es eine *komprimierte*⁹⁾ zweidimensionale Gitterpunktmenge \mathfrak{R} mit $\mathfrak{A}'(\mathfrak{R}) = \mathfrak{A}'$, $\mathfrak{A}''(\mathfrak{R}) = \mathfrak{A}''$ gibt; unter den zweidimensionalen Gitterpunkt mengen sind sonach die komprimierten Gitterpunkt mengen \mathfrak{R} die einzigen mit der Eigenschaft, daß es außer \mathfrak{R} keine zweite Gitterpunktmenge \mathfrak{M} mit $\mathfrak{A}'(\mathfrak{M}) = \mathfrak{A}'(\mathfrak{R})$, $\mathfrak{A}''(\mathfrak{M}) = \mathfrak{A}''(\mathfrak{R})$ gibt.

4. Ehe wir uns nun den entsprechenden Fragen für $n > 2$ zuwenden, wollen wir den auf den Fall $n = 2$ bezüglichen Sätzen (a), (b) eine etwas andere Form geben. Wie wir gleich sehen werden, erweisen sich diese Sätze nämlich als gleichbedeutend mit den folgenden drei Sätzen:

(A) Die Lösungsanzahl $N(\mathfrak{A}'|\mathfrak{A}'')$ ist dann und nur dann > 0 , wenn es Partitionen \mathfrak{C}' , \mathfrak{C}'' und eine 2-dimensionale komprimierte Gitterpunktmenge \mathfrak{R} gibt, für die $\mathfrak{A}'(\mathfrak{R}) = \mathfrak{C}'$, $\mathfrak{A}''(\mathfrak{R}) = \mathfrak{C}''$ und

$$(3) \quad \mathfrak{C}' \geq \mathfrak{A}', \quad \mathfrak{C}'' \geq \mathfrak{A}''$$

gilt.

(B) Wenn die Lösungsanzahl $N(\mathfrak{A}'|\mathfrak{A}'') = 1$ ist, d. h. wenn es eine einzige 2-dimensionale Gitterpunktmenge \mathfrak{M} mit $\mathfrak{A}'(\mathfrak{M}) = \mathfrak{A}'$, $\mathfrak{A}''(\mathfrak{M}) = \mathfrak{A}''$ gibt, dann ist die Gitterpunktmenge \mathfrak{M} allemal komprimiert.

(C) Wenn es eine komprimierte 2-dimensionale Gitterpunktmenge \mathfrak{R} mit $\mathfrak{A}'(\mathfrak{R}) = \mathfrak{A}'$, $\mathfrak{A}''(\mathfrak{R}) = \mathfrak{A}''$ gibt, dann ist stets $N(\mathfrak{A}'|\mathfrak{A}'') = 1$, d. h. es gibt außer \mathfrak{R} keine zweite Gitterpunktmenge \mathfrak{M} , für welche $\mathfrak{A}'(\mathfrak{M}) = \mathfrak{A}'$, $\mathfrak{A}''(\mathfrak{M}) = \mathfrak{A}''$ wäre.

Um zunächst zu sehen, daß eine notwendige und hinreichende Bedingung für $N(\mathfrak{A}'|\mathfrak{A}'') > 0$ ebensowohl durch Satz (A) wie durch Satz (a) geliefert wird, nehmen wir an, es sei $N(\mathfrak{A}'|\mathfrak{A}'') > 0$, gemäß (a) also $\mathfrak{A}'' \leq (\mathfrak{A}')^{-1}$. Wir brauchen dann nur $\mathfrak{C}' = \mathfrak{A}'$, $\mathfrak{C}'' = (\mathfrak{A}')^{-1}$ zu setzen, um zu sehen, daß die Bedingung von Satz (A) erfüllt ist, wenn für \mathfrak{R} diejenige (eindeutig bestimmte) komprimierte Menge genommen wird, zu der die konjugierten Partitionen \mathfrak{C}' und $\mathfrak{C}'' = (\mathfrak{C}')^{-1}$ gehören. Wenn umgekehrt eine komprimierte

⁹⁾ Allgemein heißt eine Menge \mathfrak{R} von Gitterpunkten des n -dimensionalen Raumes „komprimiert“, wenn zugleich mit jedem zu \mathfrak{R} gehörenden Gitterpunkt (x'_1, \dots, x'_n) auch alle den Ungleichungen $x_r \leq x'_r$ ($1 \leq r \leq n$) genügenden Gitterpunkte (x_1, \dots, x_n) zu \mathfrak{R} gehören. Siehe „Gitterpunktfiguren“ II. Komprimierte Gitterpunkt mengen. Sitz.ber. d. Bayer. Akad., Jahrg. 1940, S. 69–131, Nr. 1, 2, S. 70, „Gitterpunktfiguren“, Note V. Der Hauptsatz über den Umbau komprimierter Gitterpunkt mengen, ebenda, Jahrg. 1941, S. 1–37, Nr. 3, S. 4, „Über die Anzahl komprimierter Gitterpunkt mengen von gegebener Punktezahl“, Math. Zeitschr. 47 (1941), S. 352, sowie die in Anm. 14 genannte Arbeit in Crelles Journal, Nr. 3.

Menge \mathfrak{R} mit $\mathfrak{A}^{(\nu)}(\mathfrak{R}) = \mathfrak{C}^{(\nu)}$ ($\nu = 1, 2$) existiert, so daß (3) erfüllt ist, dann bedeutet dies, daß $\mathfrak{C}'' = (\mathfrak{C}')^{-1}$ und somit $\mathfrak{A}'' \leq (\mathfrak{C}')^{-1}$ ist. Zuzufolge (3) ist aber außerdem $(\mathfrak{C}')^{-1} \leq (\mathfrak{A}')^{-1}$, also $\mathfrak{A}'' \leq (\mathfrak{A}')^{-1}$, und gemäß (a) ist daher $N(\mathfrak{A}' | \mathfrak{A}'') \geq 0$, so daß die Bedingung von Satz (A) sich auch als hinreichend erweist.

Hinsichtlich der Sätze (B) und (C) aber ist zu sagen, daß sie sich inhaltlich decken mit dem, was in (b) in eine einzige Aussage zusammengefaßt ist.

5. Was nun den Fall von $n \geq 3$ Dimensionen betrifft¹⁰⁾, so wollen wir beweisen¹¹⁾, daß die folgenden zu (A) und (B) völlig analogen Sätze gelten¹²⁾:

Satz 1. Es gibt zu $n (\geq 2)$ Partitionen $\mathfrak{A}', \dots, \mathfrak{A}^{(n)}$ der Zahl m dann und nur dann eine n -dimensionale Gitterpunktmenge \mathfrak{R} mit $\mathfrak{A}^{(\nu)}(\mathfrak{R}) = \mathfrak{A}^{(\nu)}$ ($\nu = 1, \dots, n$), d. h. es ist dann und nur dann $N(\mathfrak{A}' | \dots | \mathfrak{A}^{(n)}) > 0$, wenn es eine n -dimensionale komprimierte Gitterpunktmenge \mathfrak{R} mit

$$(4) \quad \mathfrak{A}^{(\nu)}(\mathfrak{R}) = \mathfrak{C}^{(\nu)} \geq \mathfrak{A}^{(\nu)} \quad (1 \leq \nu \leq n)$$

gibt¹³⁾.

Satz 2. Wenn es zu $n (\geq 2)$ vorgegebenen Partitionen $\mathfrak{A}', \dots, \mathfrak{A}^{(n)}$ der Zahl m nur eine einzige n -dimensionale Gitterpunktmenge \mathfrak{R} mit $\mathfrak{A}^{(\nu)}(\mathfrak{R}) = \mathfrak{A}^{(\nu)}$ ($1 \leq \nu \leq n$) gibt, also $N(\mathfrak{A}' | \dots | \mathfrak{A}^{(n)}) = 1$ ist, dann ist diese Menge \mathfrak{R} stets komprimiert.

¹⁰⁾ Im Falle $n = 1$ liegt kein Problem vor: Für jedes $m \geq 1$ ist $N(\mathfrak{A}') > 0$ und zwar $= 1$, wenn \mathfrak{A}' gleich der Partition $(1, 1, \dots, 1)$ ist; für jede andere Partition \mathfrak{A}' von m ist $N(\mathfrak{A}') = 0$; vgl. Note I, l. c. ¹⁾, S. 32, Formeln (9).

¹¹⁾ An den Tabellen der Werte der Zahlen N , die in Note IV, l. c. ²⁾, S. 164f. für $n = 3$ und $1 \leq m \leq 5$ gegeben wurden, fiel unmittelbar auf, daß dort allemal, wo sich $N = 1$ ergibt, die zugehörige Gitterpunktmenge komprimiert ist (vgl. l. c. S. 162, Nr. 15, erster Absatz), was sich nun gemäß Satz 2 allgemein für beliebiges m und n als gültig erweist. Es ließ sich aber an diesen Tabellen auch — nach Aufstellung aller komprimierten 3-dimensionalen Gitterpunktmengen \mathfrak{R} eines bestimmten Grades m ($1 \leq m \leq 5$) und der zugehörigen Partitionensysteme $\mathfrak{C}', \mathfrak{C}'', \mathfrak{C}'''$ — feststellen, daß für keine anderen Systeme $\mathfrak{A}', \mathfrak{A}'', \mathfrak{A}'''$ sich $N(\mathfrak{A}' | \mathfrak{A}'' | \mathfrak{A}''') > 0$ ergab außer für jene, die zu einem solchen System $\mathfrak{C}', \mathfrak{C}'', \mathfrak{C}'''$ in der Beziehung (4) stehen. Daß für die letzteren Systeme $\mathfrak{A}', \mathfrak{A}'', \mathfrak{A}'''$ notwendig $N > 0$ ausfallen mußte, war aus dem Satz in Note III, l. c. ²⁾ zu ersehen. Daß aber nur für diese Systeme $N(\mathfrak{A}' | \mathfrak{A}'' | \mathfrak{A}''') > 0$ ist, wird nun durch Satz 1 allgemein für beliebiges m und n bestätigt.

¹²⁾ Bezüglich Satz 1 vgl. Sitz.ber. d. Bayer. Akad. d. Wiss., Jahrg. 1941, S. 17*, Sitzung vom 25. Oktober 1941, Punkt 3, sowie den Hinweis auf diesen Satz am Schluß der vorläufigen Mitteilung l. c. ¹³⁾.

¹³⁾ Durch diesen Satz wird die Aufstellung derjenigen Partitionensysteme $\mathfrak{A}', \dots, \mathfrak{A}^{(n)}$, für welche $N > 0$ ist, zurückgeführt auf die Aufstellung aller komprimierten Gitterpunktmengen \mathfrak{R} . Wie diese letzteren vollständig zu gewinnen sind, ergibt sich aus „Gitterpunktfiguren“, Note II, l. c. ²⁾, §§ 7, 8 (für $n = 3$) und Note V, l. c. ²⁾ § 7.

Andererseits ist bekannt, daß der zu Satz (C) analoge Satz für $n \geq 3$ Dimensionen *nicht* zutrifft, daß es vielmehr in drei und mehr Dimensionen komprimierte Gitterpunkt mengen gibt, die nicht die einzigen Gitterpunkt mengen mit den gleichen zugehörigen Partitionen sind ¹⁴⁾.

Der Beweis der Sätze 1 und 2 wird in § 3 gegeben, nachdem in § 2 die dabei verwendeten Hilfssätze über Kompression von Gitterpunkt mengen und das Verhalten der zugehörigen Partitionen bei Kompression gebracht werden.

§ 2.

Kompression von Gitterpunkt mengen.

6. Sei ν eine der Zahlen $1, \dots, n$. Eine Gitterpunktmenge ¹⁵⁾ des n -dimensionalen Raumes mit den Koordinaten x_1, \dots, x_n heiße ¹⁶⁾ „ x_ν -komprimiert“, wenn mit jedem Punkt (x_1^0, \dots, x_n^0) , der zu \mathfrak{M} gehört, stets auch alle jene Gitterpunkte (x_1, \dots, x_n) zu \mathfrak{M} gehören, für die $0 < x_\nu \leq x_\nu^0$ und $x_i = x_i^0$ ($i \neq \nu$) gilt. Für p Zahlen ν_1, \dots, ν_p aus der Reihe $1, \dots, n$ werde als „ $(x_{\nu_1} \dots x_{\nu_p})$ -komprimiert“ eine Menge bezeichnet, die x_{ν_k} -komprimiert ist für jeden der Werte $k = 1, \dots, p$. Es bedeutet dann „ $(x_1 \dots x_n)$ -komprimiert“ dasselbe wie „komprimiert“ schlechthin im Sinne der in Anm. 9 wieder gegebenen Definition.

Die eben besprochenen, teilweise (nämlich nur in gewissen Achsenrichtungen) komprimierten Gitterpunkt mengen spielen eine Rolle bei den zu besprechenden Übergängen von beliebigen Gitterpunkt mengen zu komprimierten Gitterpunkt mengen.

7. Sei ¹⁷⁾ $n \geq 2$ und \mathfrak{M} eine Gitterpunktmenge im $(x_1 \dots x_n)$ -Raum. Für jedes Wertsystem $(x') = (x'_1, \dots, x'_n)$ von $n - 1$ positiven ganzen Zahlen bedeute $m(x')$ die Anzahl der Gitterpunkte (x_1, x_2, \dots, x_n) aus \mathfrak{M} , für welche $x_2 = x'_2, \dots, x_n = x'_n$ ist; mit \mathfrak{M}^* werde diejenige Gitterpunktmenge bezeichnet, die für jedes Wertsystem (x') alle Gitterpunkte (x_1, x'_2, \dots, x'_n) mit

$$(5) \quad 0 < x_1 \leq m(x')$$

¹⁴⁾ Vgl. Note II, l. c. ⁹⁾, Nr. 6, S. 73 sowie „Komprimierte Gitterpunkt mengen und eine additivzahlentheoretische Aufgabe“ (erscheint in Crelles Journal für Mathematik, Bd. 184), wo gezeigt ist, daß zu jedem $n \geq 3$ und jeder beliebig großen Zahl G eine komprimierte Gitterpunktmenge \mathfrak{R} gefunden werden kann, so daß für die Partitionen $\mathfrak{Q}^{(\nu)} = \mathfrak{Q}^{(\nu)}(\mathfrak{R})$ die Anzahl $N(\mathfrak{Q}' | \dots | \mathfrak{Q}^{(n)}) > G$ ausfällt.

¹⁵⁾ Die Endlichkeit dieser Menge braucht nicht vorausgesetzt zu werden. Erst später (von Hilfssatz 2 in Nr. 8 an) werden wir uns auf endliche Gitterpunkt mengen beschränken.

¹⁶⁾ Vgl. Note VII, l. c. ²¹⁾, Nr. 2.

¹⁷⁾ Wegen $n = 1$ siehe Anm. 20.

umfaßt¹⁸⁾. Als „ x_1 -Kompression“ von \mathfrak{M} werde der Übergang von \mathfrak{M} zu \mathfrak{M}^* bezeichnet, wobei also, wie man sagen kann, auf jeder zur x_1 -Achse parallelen Geraden die Gitterpunkte in Richtung abnehmender x_1 -Werte zusammen-geschoben werden¹⁹⁾. Die Menge \mathfrak{M}^* ist natürlich x_1 -komprimiert und es ist $\mathfrak{M}^* = \mathfrak{M}$, wenn \mathfrak{M} bereits selbst x_1 -komprimiert ist.

Analog wird die x_r -Kompression für jedes $r \leq n$ definiert²⁰⁾.

8. Über die besprochenen Änderungen von Gitterpunktmengen durch Kompression in einer Achsenrichtung gelten dann die folgenden an anderer Stelle bewiesenen Sätze:

Hilfssatz 1. Aus einer $(x_1 \dots x_q)$ -komprimierten Gitterpunktmenge des $(x_1 \dots x_n)$ -Raumes (wobei $1 \leq q < n$ sei) entsteht durch (x_{q+1}) -Kompression eine Menge, die wieder $(x_1 \dots x_q)$ -komprimiert, insgesamt also $(x_1 \dots x_q x_{q+1})$ -komprimiert ist²¹⁾.

Hilfssatz 2. Sei \mathfrak{M} eine endliche Gitterpunktmenge des $(x_1 \dots x_n)$ -Raumes, \mathfrak{M}^* die aus \mathfrak{M} durch x_1 -Kompression entstehende Menge, ferner \mathfrak{P}' bzw. \mathfrak{P}'_* die zur Menge \mathfrak{M} bzw. \mathfrak{M}^* gehörende auf die x_1 -Koordinate bezügliche Partition²²⁾; sei ferner \mathfrak{M} so beschaffen, daß die Zahlen a'_μ der Partition $\mathfrak{P}' = (a'_1, a'_2, \dots)$ der Ungleichung $a'_1 \geq a'_2 \geq \dots$ genügen. Dann gilt für die Partitionen \mathfrak{P}' und \mathfrak{P}'_* die Ordnungsbeziehung

$$\mathfrak{P}'_* \geq \mathfrak{P}',$$

wobei nur dann $\mathfrak{P}'_* = \mathfrak{P}'$ ist, wenn \mathfrak{M} eine x_1 -komprimierte Menge, somit $\mathfrak{M}^* = \mathfrak{M}$ ist. Analog ist natürlich allgemein $\mathfrak{P}^{(r)}_* \geq \mathfrak{P}^{(r)}$, wenn \mathfrak{M}^* aus \mathfrak{M} durch x_r -Kompression entsteht²³⁾.

¹⁸⁾ Bei einer unendlichen Gitterpunktmenge \mathfrak{M} (vgl. Anm. 15) kann natürlich auch $m(x) = \infty$ sein und es sind dann in \mathfrak{M}^* alle Punkte mit $x_r = x'_r$ ($2 \leq r \leq n$) aufzunehmen.

¹⁹⁾ Vgl. die vorläufige Mitteilung in Sitz.ber. d. Bayer. Akad. d. Wiss., Jahrg. 1941, S. 1*, Sitzung vom 11. Januar, Punkt 2; ferner Note VII, l. c. ²¹⁾, Nr. 3.

²⁰⁾ Natürlich kann man die gegebene Definition auch auf den Fall $n = 1$ ausdehnen; wieder wird \mathfrak{M}^* durch die Ungleichung (5) erklärt, wobei $m(x)$ zu ersetzen ist durch die Anzahl m der Gitterpunkte von \mathfrak{M} .

²¹⁾ Vgl. „Gitterpunktfiguren“ VII. Schrittweise Kompression partiell-komprimierter Mengen. Sitz.ber. d. Bayer. Akad. d. Wiss., Jahrg. 1941, S. 165–170, Nr. 4, Satz 1.

²²⁾ Sind $\mathfrak{P}'', \dots, \mathfrak{P}^{(n)}$ bzw. $\mathfrak{P}''_*, \dots, \mathfrak{P}^{(n)}_*$ die auf die übrigen Koordinaten bezüglichen zu \mathfrak{M} bzw. \mathfrak{M}^* gehörenden Partitionen, so ist natürlich $\mathfrak{P}^{(r)}_* = \mathfrak{P}^{(r)}$ für $2 \leq r \leq n$.

²³⁾ Vgl. „Gitterpunktfiguren“ VIII. Auswirkung der Kompression von Gitterpunktmengen auf die zugehörigen Partitionen. Sitz.ber. d. Bayer. Akad. d. Wiss. 1941. S. 171–186, § 1, Nr. 2, Satz 1.

§ 3.

Beweis der Sätze 1 und 2.

9. Sei \mathfrak{M} eine endliche Gitterpunktmenge des $(x_1 \dots x_n)$ -Raumes, m die Anzahl ihrer Punkte und seien $\mathfrak{A}', \dots, \mathfrak{A}^{(n)}$ die zu \mathfrak{M} gehörenden Partitionen der Zahl m , wobei (vgl. Nr. 1) $a_1^{(1)} \geq a_2^{(1)} \geq \dots$ angenommen werden möge. Mit

$$(6) \quad \mathfrak{M}_0 = \mathfrak{M}, \quad \mathfrak{M}_1, \dots, \quad \mathfrak{M}_n = \mathfrak{A}$$

bezeichnen wir diejenigen Mengen, von denen jeweils \mathfrak{M}_ν durch x_ν -Kompression aus $\mathfrak{M}_{\nu-1}$ entsteht ($1 \leq \nu \leq n$). Dem Hilfssatz 1 (Nr. 8) zufolge ist dann jede Menge \mathfrak{M}_ν eine $(x_1 \dots x_\nu)$ -komprimierte Menge, es ist somit $\mathfrak{M}_n = \mathfrak{A}$ komprimiert. Gemäß Hilfssatz 2 ändert sich aber beim Übergang von $\mathfrak{M}_{\nu-1}$ zu \mathfrak{M}_ν von den zugehörigen Partitionen höchstens die eine auf die x_ν -Koordinate bezügliche; und wenn sie sich ändert, dann stets so, daß zu \mathfrak{M}_ν eine „höhere“ Partition gehört als zu $\mathfrak{M}_{\nu-1}$. Für die zu \mathfrak{A} gehörenden Partitionen folgen daraus die Beziehungen (4), womit Satz 1 von Nr. 5 bewiesen ist.

10. Andererseits gilt für die Lösungsanzahlen $N(\mathfrak{A}' | \mathfrak{A}'' | \dots | \mathfrak{A}^{(n)})$ der folgende Satz:

Hilfssatz 3. Wenn zwischen den $2n$ Partitionen $\mathfrak{A}', \dots, \mathfrak{A}^{(n)}, \mathfrak{C}', \dots, \mathfrak{C}^{(n)}$ derselben Zahl m die Beziehungen bestehen

$$(7) \quad \mathfrak{C}' \geq \mathfrak{A}', \dots, \quad \mathfrak{C}^{(n)} \geq \mathfrak{A}^{(n)},$$

wenn ferner $N(\mathfrak{C}' | \dots | \mathfrak{C}^{(n)}) \neq 0$ (also ≥ 1) ist, dann ist

$$(8) \quad N(\mathfrak{A}' | \dots | \mathfrak{A}^{(n)}) \geq N(\mathfrak{C}' | \dots | \mathfrak{C}^{(n)})$$

und es gilt in (8) nur dann das Gleichheitszeichen, wenn es in allen n Relationen (7) gilt²⁴⁾.

Unter den Voraussetzungen des Satzes 2 in Nr. 5 gelten nun für die aus \mathfrak{M} gemäß Nr. 9 gebildete komprimierte Menge $\mathfrak{M}_n = \mathfrak{A}$ und die zu \mathfrak{A} gehörenden Partitionen $\mathfrak{C}', \dots, \mathfrak{C}^{(n)}$ gemäß Satz 1 die Beziehungen (4); dabei ist

$$(9) \quad N(\mathfrak{C}' | \dots | \mathfrak{C}^{(n)}) \geq 1,$$

da ja mindestens die eine Menge \mathfrak{A} existiert, für welche $\mathfrak{C}', \dots, \mathfrak{C}^{(n)}$ die zugehörigen Partitionen sind. Nach Hilfssatz 3 gilt also (8), wobei — wegen der Voraussetzung $N(\mathfrak{A}' | \dots | \mathfrak{A}^{(n)}) = 1$ — in (8) und daher in allen Relationen (4) notwendig das Gleichheitszeichen gilt. Das ist aber, wie aus Hilfssatz 2 ersichtlich, nur so möglich, daß in der Folge (6) der aus \mathfrak{M} durch schrittweise

²⁴⁾ Dieser Satz ist eine unmittelbare Folge des in Note III, l. c. ²⁾ bewiesenen Satzes.

Kompression gebildeten Mengen durchweg $\mathfrak{M}_{-1} = \mathfrak{M}_1$, also $\mathfrak{M}_n = \mathfrak{M}_0$, das heißt $\mathfrak{R} = \mathfrak{M}$ ist. Die Menge \mathfrak{M} ist also komprimiert. Satz 2 ist damit bewiesen.

Dafür, daß eine Gitterpunktmenge \mathfrak{M} bezüglich des Systems der zugehörigen Partitionen einzig ist, gibt natürlich Satz 2 nur eine notwendige Bedingung: die Komprimiertheit von \mathfrak{M} . Daß diese Bedingung (für $n \geq 3$) nicht hinreicht, geht aus den Bemerkungen in Nr. 5 (siehe insbesondere Anm. 14) über die Ungültigkeit des zu Satz (C) analogen Satzes für $n > 2$ hervor. Satz 2 stellt also zu der eingangs (am Schluß von Nr. 1) bezüglich $N = 1$ gestellten Frage keineswegs eine volle Beantwortung dar, wie sie durch Satz 1 (Nr. 5) für die auf $N > 0$ bezügliche Frage gegeben wird.

(Eingegangen am 9. 11. 1941.)

Schnittpunktzahl rektifizierbarer und nichtrektifizierbarer Kurven.

Von

Wilhelm Maak in Hamburg.

Wenn \mathcal{K} eine feste Kurve und \mathcal{K}' eine bewegliche Kurve in der Ebene ist, so wird sich die Anzahl ihrer Schnittpunkte ändern mit der Lage von \mathcal{K}' . Es ist möglich, eine mittlere Schnittpunktzahl mit Hilfe eines Integrals zu erklären, und es gilt der Satz: *Die mittlere Schnittpunktzahl ist gleich dem Produkt der Kurvenlängen (Poincaréformel).* In einer früheren Arbeit habe ich diese Formel für rektifizierbare Kurven bewiesen¹⁾. Es war mir aber nicht möglich, mit der damals angewandten Methode zu zeigen, daß das betreffende Integral unendlich wird, wenn beide Kurven nicht rektifizierbar sind, also unendliche Länge haben.

Durch eine bessere Beweisanordnung ist es mir gelungen, die Herleitung der Poincaréformel erheblich zu vereinfachen, vor allem aber, sie so zu gestalten, daß man ohne weiteres erkennt, daß die mittlere Schnittpunktzahl unendlich wird, wenn eine oder beide Kurven unendliche Länge besitzen. Damit ist die Poincaréformel also für beliebige einfache, stetige Kurven bewiesen.

Aus dieser allgemeinen Formel kann man sofort das folgende interessante Kriterium für Rektifizierbarkeit von Kurven ablesen: *Haben zwei Kurven bei jeder gegenseitigen Lage eine beschränkte Anzahl Schnittpunkte, so sind beide Kurven rektifizierbar.*

§ 1.

Einleitung.

Unter einer Kurve \mathcal{K} versteht man ein eindeutiges und stetiges Abbild der Punkte t der Einheitsstrecke $0 \leq t \leq 1$ auf Punkte der Euklidischen Ebene. Wir bezeichnen die Abbildungsfunktion mit $\mathbf{x}(t)$. Wenn $\mathbf{x}(t)$ umkehrbar eindeutig ist, so heißt die Kurve *einfach*, wenn $\mathbf{x}(0) = \mathbf{x}(1)$ ist, so nennen wir sie *geschlossen*.

Auf \mathcal{K} geben wir nun irgendwelche endlich viele Punkte $t_0 = 0 < t_1 < \dots < t_{r-1} < t_r = 1$ vor. Wir verbinden sie durch r Strecken s_i , ihre Länge sei s_i . So erhalten wir einen „einschriebenen“ Streckenzug \mathfrak{S} , oder

¹⁾ W. Maak, Über stetige Kurven. Abh. Math. Sem. Hans. Univers. 12, S. 163.

auch $\mathfrak{S}^{(d)}$, wenn für alle i gilt $|t_i - t_{i-1}| < \delta$. Die Länge von \mathfrak{S} oder $\mathfrak{S}^{(d)}$ bezeichnen wir mit S bzw. $S^{(d)}$. Die Länge L von \mathfrak{R} ist dann erklärt durch

$$L = \text{ob. Gr. } S,$$

wobei die obere Grenze bezüglich aller einbeschriebenen Streckenzüge gemeint ist. Es gilt

$$L = \lim_{\delta \rightarrow 0} S^{(d)}.$$

Wenn L endlich ist, so heißt \mathfrak{R} *rektifizierbar* oder *streckbar*.

Ist \mathfrak{R}' eine andere Kurve, die wir als einfach annehmen wollen, so nenne ich einen Punkt aus dem Durchschnitt $\mathfrak{D}(\mathfrak{R}, \mathfrak{R}')$ von \mathfrak{R} und \mathfrak{R}' dann und nur dann einen *Schnittpunkt*, wenn an dieser Stelle \mathfrak{R} tatsächlich von einem Ufer von \mathfrak{R}' zum anderen übertritt. Unter Anzahl N von Schnittpunkten von \mathfrak{R} und \mathfrak{R}' verstehe ich die Anzahl der t , für die $\mathfrak{x}(t)$ Schnittpunkt ist. Wenn \mathfrak{R}' starr beweglich ist, so ist N eine Funktion von \mathfrak{R}' , wir schreiben

$$(1) \quad N = N(\mathfrak{R}').$$

Die Lage einer starr beweglichen Kurve \mathfrak{R}' kann man angeben, indem man mit ihr ein Linienelement starr verbindet, das ist ein Punkt (x_1, x_2) mit einer Richtung φ . Jeder Lage von \mathfrak{R}' entspricht so ein Punkt (x_1, x_2, φ) im dreidimensionalen Raume. Auf diese Weise wird aus (1) eine Funktion von drei Veränderlichen

$$N = N(\mathfrak{R}') = N(x_1, x_2, \varphi).$$

Wie alle anderen in dieser Arbeit auftretenden Funktionen, so ist auch diese meßbar¹⁾.

Es sei $f(\mathfrak{R}') = f(x_1, x_2, \varphi)$ irgendeine positive meßbare Funktion, dann führt man für ihr Integral folgende Schreibweise ein

$$\iiint f(x_1, x_2, \varphi) dx_1 dx_2 d\varphi = \int f \dot{\mathfrak{R}}'.$$

Man nennt

$$\dot{\mathfrak{R}}' = dx_1 dx_2 d\varphi$$

die kinematische Dichte von $\mathfrak{R}'^2)$. Sie hat einige wichtige Eigenschaften.

1. Die kinematische Dichte ist *bewegungsinvariant*. Das bedeutet z. B., daß das Integral von (1)

$$(2) \quad \int N \dot{\mathfrak{R}}',$$

welches ich als *mittlere Schnittpunktzahl* von \mathfrak{R} und \mathfrak{R}' bezeichne, unabhängig von der Lage von \mathfrak{R} ist.

²⁾ W. Blaschke, Vorlesungen über Integralgeometrie. 1. Heft. Zweite Auflage. Leipzig und Berlin 1936.

2. Die kinematische Dichte ist *wahlinvariant*. Für (2) bedeutet das, daß dies Integral nicht davon abhängt, welch Linienelement ich zur Beschreibung der Lage von \mathcal{R}' benutze.

3. Die kinematische Dichte ist *umkehrinvariant*. Dies bedeutet für (2), daß das Integral sich nicht ändert, wenn ich statt \mathcal{R} das \mathcal{R}' festhalte und \mathcal{R} als beweglich anschau.

Das wichtigste Hilfsmittel beim Beweise der Poincaréformel ist Lebesgues Satz vom gliedweisen Integrieren: Es seien die Funktionen $f_i = f_i(\mathcal{R}')$ meßbar und $g = g(\mathcal{R}')$ sei integrierbar. Außerdem gelte

$$(3) \quad \begin{aligned} 0 &\leq f_i \leq g, \\ f &= \lim_{i \rightarrow \infty} f_i; \end{aligned}$$

dann darf (3) gliedweise integriert werden:

$$\int f \mathcal{R}' = \lim_{i \rightarrow \infty} \int f_i \mathcal{R}'.$$

Für manche Zwecke ist folgende Fassung des Theorems zweckmäßiger: Die Integrale der summierbaren Funktionen f_i seien gleichmäßig beschränkt

$$\int f_i \mathcal{R}' \leq M$$

und es gelte

$$(4) \quad \begin{aligned} \lim_{i \rightarrow \infty} f_i &= f, \\ 0 &\leq f_i \leq f, \end{aligned}$$

dann darf man (4) gliedweise integrieren.

§ 2.

Die Poincaréformel.

Es seien wie im § 1 \mathcal{R} und \mathcal{R}' Kurven, \mathcal{R} sei fest und \mathcal{R}' beweglich, ihre Schnittpunktzahl heiße $N(\mathcal{R}')$. Die zu beweisende Formel lautet dann³⁾

$$\int N \mathcal{R}' = 4LL'.$$

Diese Formel ist für Streckenzüge so leicht herzuleiten, daß ich auf diesen Fall nicht einzugehen brauche.

Zunächst sollen beide Kurven einfach sein und \mathcal{R}' zudem geschlossen. Ein der Kurve \mathcal{R} einbeschriebener Streckenzug $\mathfrak{S}^{(d)}$ habe die Strecken $s_i^{(d)}$. Dann erkläre ich folgende Funktion von \mathcal{R}' :

$$p_i^{(d)} = \begin{cases} 1, & \text{wenn der eine Endpunkt von } s_i^{(d)} \text{ im Innern von } \mathcal{R}' \text{ liegt, der} \\ & \text{andere außen;} \\ 0, & \text{wenn beide Endpunkte auf derselben Seite von } \mathcal{R}' \text{ liegen.} \end{cases}$$

³⁾ Der Faktor 4 wurde am Anfang der Arbeit der Einfachheit halber nicht erwähnt.

Wenn also $p_i^{(\delta)} = 1$, so schneidet $s_i^{(\delta)}$ und auch das dem $s_i^{(\delta)}$ entsprechende Stück von \mathcal{R} die Kurve \mathcal{R}' mindestens einmal. Deshalb gilt

$$(1) \quad \sum_i \dot{p}_i^{(\delta)} \leq N.$$

Sorge ich dafür, daß bei kleiner werdendem δ die Eckpunkte von $\mathfrak{S}^{(\delta)}$ an immer neuen Punkten von \mathcal{R} liegen, so gilt auch

$$(2) \quad \lim_{\delta \rightarrow 0} \sum_i p_i^{(\delta)} = N.$$

Mit $p_i^{(\delta)}$ ist auch $\sum_i p_i^{(\delta)}$ eine Funktion von \mathcal{R}' . Ich suche ihr Integral zu berechnen:

$$(3) \quad \int \sum_i p_i^{(\delta)} \dot{\mathcal{R}}' = \sum_i \int p_i^{(\delta)} \dot{\mathcal{R}}'.$$

Dazu „strecke“ ich jetzt $\mathfrak{S}^{(\delta)}$, das heißt ich trage sämtliche Strecken $s_i^{(\delta)}$ auf einer festen Geraden g von einem festen Punkt P aus hintereinander ab. Dabei ändern sich wegen der Bewegungsinvarianz des kinematischen Maßes (bzw. der entsprechenden Dichte) die einzelnen Integrale auf der rechten Seite von (3) nicht. Die neuen Funktionen $p_i^{(\delta)}$ will ich der Klarheit wegen $\pi_i^{(\delta)}$ nennen. Also

$$(4) \quad \int \sum_i p_i^{(\delta)} \dot{\mathcal{R}}' = \sum_i \int p_i^{(\delta)} \dot{\mathcal{R}}' = \sum_i \int \pi_i^{(\delta)} \dot{\mathcal{R}}' = \int \sum_i \pi_i^{(\delta)} \dot{\mathcal{R}}'.$$

Jetzt setze ich vorläufig \mathcal{R} als streckbar voraus von der Länge L , und s soll eine Strecke von derselben Länge L sein, die ich ebenfalls von P aus auf g abtrage. Es habe s mit \mathcal{R}' etwa n Schnittpunkte. Dann gilt

$$(5) \quad \begin{aligned} \sum_i \pi_i^{(\delta)} &\leq n, \\ \lim_{\delta \rightarrow 0} \sum_i \pi_i^{(\delta)} &= n. \end{aligned}$$

Diese Gleichung kann ich jedenfalls erzwingen, indem ich Sorge trage, daß die Endpunkte der Strecken, die auf g abgetragen werden, an immer neuen Stellen liegen, wenn δ gegen Null geht. Ist nun n integrierbar, so darf ich wegen (5) gliedweise integrieren und man erhält

$$(6) \quad \lim_{\delta \rightarrow 0} \int \sum_i \pi_i^{(\delta)} \dot{\mathcal{R}}' = \int n \dot{\mathcal{R}}'.$$

Außerdem ergibt sich aus (4) und (5)

$$\int \sum_i p_i^{(\delta)} \dot{\mathcal{R}}' \leq \int n \dot{\mathcal{R}}'.$$

Hieraus und aus (1) und (2) folgt (ebenfalls nach Lebesgues)

$$(7) \quad \lim_{\delta \rightarrow 0} \int \sum_i p_i^{(\delta)} \dot{\mathcal{R}}' = \int N \dot{\mathcal{R}}'.$$

Aus (4), (6) und (7) entnehme ich

$$(8) \quad \int N \mathfrak{K}' = \int n \mathfrak{K}'.$$

Diese Formel gilt speziell auch dann, wenn \mathfrak{K}' ein gleichseitiges Dreieck ist. In diesem Falle ist n stets integabel, weil wir für Streckenzüge die Poincaréformel schon als bekannt voraussetzen können. Es braucht also n nicht ausdrücklich als integabel vorausgesetzt zu werden. Daraus folgt sofort, daß (8) auch dann richtig sein muß, wenn \mathfrak{K}' eine Strecke ist, nämlich etwa eine der drei Seiten des Dreiecks. Hat diese Strecke die Länge l , so ergibt sich

$$\int n \mathfrak{K}' = 4 L l.$$

Also

$$(9) \quad \int N \mathfrak{K}' = 4 L l.$$

Das ist die Poincaréformel für eine feste streckbare Kurve und eine bewegliche Strecke.

Es ist aber das kinematische Maß umkehrinvariant. Deshalb muß die Formel (9) auch dann gelten, wenn die Strecke fest ist und die streckbare Kurve beweglich. Dieser Fall liegt nun gerade bei (8) vor, wenn ich nur jetzt auch \mathfrak{K}' als streckbar annehme. Dann ist also wirklich, auch ohne daß ich es ausdrücklich voraussetze, n integabel und (8) bewiesen. Aus (8) und (9) folgt, weil s die Länge L und \mathfrak{K}' die Länge L' hat,

$$\int N \mathfrak{K}' = 4 L L'.$$

Dies ist die Poincaréformel für zwei streckbare einfache Kurven, von denen eine auch noch geschlossen ist. Nachträglich kann man ohne Schwierigkeit zeigen, daß die eine Kurve auch nicht geschlossen zu sein braucht, und daß Doppelpunkte bis zu einem gewissen Grade zugelassen werden können¹⁾.

Es bleibt jetzt noch zu untersuchen, was geschieht, wenn eine der Kurven \mathfrak{K} und \mathfrak{K}' oder beide nicht rektifizierbar sind. Es sei etwa \mathfrak{K} von unendlicher Länge und \mathfrak{K}' wieder geschlossen. Ich führe dann die Annahme, N sei integabel, zu einem Widerspruch. Wegen (1) könnte ich dann nämlich (2) gliedweise integrieren und erhielte

$$(10) \quad \lim_{\delta \rightarrow 0} \int \sum_i p_i^{(\delta)} \mathfrak{K}' = \int N \mathfrak{K}'.$$

Es soll jetzt die Strecke s ganz beliebige Länge haben. Ich trage sie wie früher auf g von P aus ab. Für hinreichend kleine δ gilt dann

$$\sum_i \pi_i^{(\delta)} \geq n,$$

wenn n die Anzahl der Schnittpunkte von s mit \mathfrak{K}' bedeutet. Wegen (4) heißt das

$$\int \sum_i p_i^{(\delta)} \mathfrak{K}' \geq \int n \mathfrak{K}',$$

wieder für hinreichend kleine δ . Aus (10) folgt so

$$\int N \dot{R}' \geq \int n \dot{R}'.$$

Nun wird das Integral rechts sicher beliebig groß, weil ja s beliebige Länge hat. Also ist

$$\int N \dot{R}' = \infty$$

und somit N nicht integrierbar, im Gegensatz zu unserer Annahme. Wir hatten wieder \dot{R}' als geschlossen vorausgesetzt. Aber auch in diesem Falle macht der Übergang zu nicht geschlossenen Kurven keine Schwierigkeit.

Wir haben jetzt also bewiesen:

Poincaréformel: Für einfache rektifizierbare oder nicht rektifizierbare Kurven gilt

$$\int N \dot{R}' = 4 LL',$$

wo bei nicht rektifizierbaren Kurven die Länge unendlich eingesetzt werden muß.

Diesem Satze entnehmen wir folgendes

Kriterium für Rektifizierbarkeit: Zwei Kurven sind dann und nur dann rektifizierbar, wenn ihre mittlere Schnittpunktzahl endlich ist. Das ist z. B. immer der Fall, wenn die Schnittpunktzahl beschränkt ist.

(Eingegangen am 7. 10. 1941.)

Über die Bedingungen, daß eine binäre Form n -ten Grades eine n -te Potenz ist, und über die rationale Kurve n -ter Ordnung im R_n .

Von

Oskar Perron in München.

§ 1.

Zwei Fragen.

Sei n eine natürliche Zahl ≥ 2 . Die binäre Form von X, Y

$$x_0 X^n + \binom{n}{1} x_1 X^{n-1} Y + \binom{n}{2} x_2 X^{n-2} Y^2 + \dots + x_n Y^n$$

ist dann und nur dann eine n -te Potenz $(\varrho X + \sigma Y)^n$, wenn ihre Koeffizienten x_i die folgende Parameterdarstellung haben:

$$(1) \quad x_0 = \varrho^n, \quad x_1 = \varrho^{n-1} \sigma, \quad x_2 = \varrho^{n-2} \sigma^2, \quad \dots, \quad x_n = \sigma^n.$$

Dafür ist offenbar notwendig und hinreichend, daß die Determinanten der Matrix

$$(2) \quad \begin{vmatrix} x_0 & x_1 & \dots & x_{n-1} \\ x_1 & x_2 & \dots & x_n \end{vmatrix}$$

sämtlich verschwinden. Das sind $\frac{1}{2} n(n-1)$ Bedingungsgleichungen zwischen den Koeffizienten x_i . Man sieht aber leicht, daß bereits das Erfülltsein der n Gleichungen

$$(3) \quad \begin{vmatrix} x_0 & x_1 \\ x_1 & x_2 \end{vmatrix} = 0, \quad \begin{vmatrix} x_1 & x_2 \\ x_2 & x_3 \end{vmatrix} = 0, \quad \dots, \quad \begin{vmatrix} x_{n-2} & x_{n-1} \\ x_{n-1} & x_n \end{vmatrix} = 0, \quad \begin{vmatrix} x_{n-1} & x_0 \\ x_n & x_1 \end{vmatrix} = 0$$

genügt (also das Verschwinden der übrigen Determinanten nach sich zieht). Denn für eine Lösung von (3), bei der $x_0 = 0$ ist, folgt aus den ersten $n-1$ Gleichungen (3) der Reihe nach: $x_1 = 0, x_2 = 0, \dots, x_{n-1} = 0$, so daß diese Lösung die Gestalt (1) hat (mit $\varrho = 0$). Bei einer Lösung, bei der $x_0 \neq 0$ ist, kann man, weil es auf einen Proportionalitätsfaktor nicht ankommt, $x_0 = 1$ annehmen. Wenn dann $x_1 \neq 0$ ist und $x_1 = \sigma$ gesetzt wird, so folgt aus den ersten $n-1$ Gleichungen (3) der Reihe nach: $x_2 = \sigma^2, x_3 = \sigma^3, \dots, x_n = \sigma^n$, so daß die Lösung wiederum die Gestalt (1) hat. Wenn aber $x_1 = 0$, so folgt aus den letzten $n-1$ Gleichungen (3) von hinten angefangen der Reihe nach: $x_n = 0, x_{n-1} = 0, \dots, x_2 = 0$, so daß die Lösung abermals die Gestalt (1) hat (mit $\sigma = 0$). Die n Gleichungen (3) haben also keine anderen Lösungen als die durch (1) dargestellten.

Deutet man die z , als projektive Koordinaten im R_n , so stellen die Gleichungen (1) die C_n im R_n dar, d. h. die rationale Kurve n -ter Ordnung im R_n , die nicht in einem R_{n-1} liegt [alle Kurven dieser Art, die sogenannten rationalen Normalkurven des R_n sind kollineare Bilder voneinander und werden bei geeigneter Wahl des Koordinatensystems durch die Gleichungen (1) dargestellt]. Wir haben daher den

Satz 1. Die C_n im R_n läßt sich als Durchschnitt von n Flächen¹⁾ zweiter Ordnung darstellen.

Nun kann man fragen, ob es nicht mit weniger als n Flächen geht. Mit weniger als $n - 1$ kann es offenbar nicht gehen, weil der Durchschnitt von höchstens $n - 2$ Flächen im R_n von höherer als erster Dimension ist, also keine Kurve darstellt. Somit bleiben aber noch die folgenden zwei Fragen offen:

Frage 1. Läßt sich die C_n im R_n als Durchschnitt von $n - 1$ algebraischen Flächen darstellen?

Frage 2. Läßt sich die C_n im R_n als Durchschnitt von $n - 1$ Flächen zweiter Ordnung darstellen?

Die Antwort auf Frage 1 lautet, wie in § 2 gezeigt wird, ja.

Die Frage 2 ist für $n = 3$ bekanntlich zu verneinen. Es wäre aber falsch, zu glauben, daß sie dann für $n > 3$ erst recht zu verneinen ist. Die Antwort auf Frage 2 lautet vielmehr, wie in § 3 gezeigt wird:

ja, wenn n eine Potenz von 2 ist,

nein für alle anderen n .

Bemerkung. Ob die Antworten ebenso lauten, wenn man „die übliche Definition des reinen Punktes zugrunde legt“ (vgl. Deutsche Mathematik, Jahrg. 6, S. 2), muß ich allerdings Herrn Leidheuser entscheiden lassen, da diese Definition wohl sein Geheimnis und mir gänzlich unbekannt ist. Meinen Ausführungen liegt aber auch nicht „die Perronsche Definition des fremden Punktes“ (ibid.) zugrunde, die es gar nicht gibt. Es handelt sich vielmehr um genau dieselbe ganz eindeutige Fragestellung wie in meiner Arbeit im Band 47 der Math. Zeitschrift und wie in den dort zitierten Arbeiten von Kronecker und (trotz Herrn Leidheusers abweichendem Interpretationsversuch) von Herrn Vahlen, der ja nicht von so nebelhaften Dingen wie „reinen Punkten“, sondern von „vollständiger“ oder „isolierter“ Darstellung spricht und mit vorbildlich klaren Worten das umreißt, was damit gemeint ist, daß nämlich die zur Darstellung der Kurve benutzten Flächen keinen „außerhalb der Kurve gelegenen Schnittpunkt“ haben dürfen. Ich habe mich der Kürze halber der in der Mengenlehre inzwischen üblich gewordenen

¹⁾ Unter „Flächen“ im R_n sind in dieser Arbeit natürlich $(n - 1)$ -dimensionale Gebilde zu verstehen, also definiert durch eine homogene Gleichung zwischen den $n + 1$ Koordinaten.

Terminologie bedient und einfach das Wort „Durchschnitt“ gebraucht ohne den Zusatz „vollständig“.

§ 2.

Antwort auf Frage 1.

Satz 2. Die Kurve (1) ist der Durchschnitt der $n - 1$ Flächen

$$(4) \quad \begin{vmatrix} x_0 & x_1 & \dots & x_k \\ x_1 & x_2 & \dots & x_{k+1} \\ \dots & \dots & \dots & \dots \\ x_k & x_{k+1} & \dots & x_{2k} \end{vmatrix} = 0 \quad (k = 1, 2, \dots, n - 1),$$

wobei die x_v für $v > n$ durch 0 zu ersetzen sind²⁾.

Beweis. Daß die Gleichungen (4) durch die Werte (1) befriedigt werden, ist klar, weil beim Einsetzen von (1) in (4) bereits alle Unterdeterminanten der ersten zwei Zeilen verschwinden. Es ist noch zu zeigen, daß auch umgekehrt jede Lösung von (4) die Gestalt (1) hat. Betrachtet man nun zuerst eine Lösung von (4), bei der $x_0 = 0$ ist, so folgt aus den $n - 1$ Gleichungen (4) der Reihe nach: $x_1 = 0, \dots, x_{n-1} = 0$; die Lösung hat daher die Gestalt (1) (mit $\sigma = 0$). Betrachtet man aber eine Lösung von (4), bei der $x_0 \neq 0$ ist, so kann man $x_0 = 1$ annehmen. Setzt man dann $x_1 = \sigma$, so geht die erste Gleichung (4) über in

$$\begin{vmatrix} 1 & \sigma \\ \sigma & x_2 \end{vmatrix} = 0,$$

also ist $x_2 = \sigma^2$. Nimmt man aber an, aus den ersten $k - 1$ Gleichungen (4) hätte man bereits der Reihe nach folgern können: $x_2 = \sigma^2, \dots, x_k = \sigma^k$, so geht die k -te Gleichung (4) durch Einsetzen dieser Werte über in

$$\begin{vmatrix} 1 & \sigma & \dots & \sigma^{k-1} & \sigma^k \\ \sigma & \sigma^2 & \dots & \sigma^k & x_{k+1} \\ \dots & \dots & \dots & \dots & \dots \\ \sigma^k & x_{k+1} & \dots & x_{2k-1} & x_{2k} \end{vmatrix} = 0.$$

Zieht man hier, von hinten anfangend, von jeder Spalte die mit σ multiplizierte vorausgehende ab, so sieht man sofort, daß die Determinante gleich

$$\pm (x_{k+1} - \sigma^{k+1})^k$$

ist. Somit ist auch $x_{k+1} = \sigma^{k+1}$. Die Lösung lautet daher schließlich

$$x_0 = 1, x_1 = \sigma, \dots, x_n = \sigma^n;$$

sie hat also wieder die Gestalt (1).

²⁾ Man könnte sie ebensogut durch irgendwelche Linearformen von x_0, \dots, x_n ersetzen. Die Flächen wären dann zwar andere, aber ihr Durchschnitt wäre derselbe.

§ 3.

Antwort auf Frage 2.

Satz 3. Die Kurve (1) läßt sich als Durchschnitt von $n - 1$ Flächen zweiter Ordnung dann und nur dann darstellen, wenn n eine Potenz von 2 ist.

Beweis. Für $n = 2$ ist die Darstellbarkeit trivial; die Kurve ist dann einfach der Kegelschnitt $x_0 x_2 - x_1^2 = 0$ (eine Gleichung).

Nun nehmen wir an, für $n = m$, wo m eine gewisse Potenz von 2 ist, sei bereits eine Darstellung durch $m - 1$ Flächen zweiter Ordnung gelungen:

$$(5) \quad f_1(x_0, x_1, \dots, x_m) = 0, \dots, f_{m-1}(x_0, x_1, \dots, x_m) = 0,$$

und wollen zeigen, daß es dann auch für $n = 2m$ eine Darstellung durch $2m - 1$ Flächen zweiter Ordnung gibt. Es sind das einfach die folgenden $2m - 1$ Flächen:

$$(a) \quad f_1(x_0, x_1, \dots, x_m) = 0, \dots, f_{m-1}(x_0, x_1, \dots, x_m) = 0,$$

$$(b) \quad f_1(x_m, x_{m+1}, \dots, x_{2m}) = 0, \dots, f_{m-1}(x_m, x_{m+1}, \dots, x_{2m}) = 0,$$

$$(c) \quad x_0 x_{2m} - \binom{m}{1} x_1 x_{2m-1} + \binom{m}{2} x_2 x_{2m-2} - \dots + \binom{m}{m} x_m x_m = 0.$$

In der Tat ist zunächst klar, daß das Gleichungssystem (a), (b), (c) jedenfalls die Lösung

$$(6) \quad \begin{aligned} x_0 &= \varrho^{2m}, \quad x_1 = \varrho^{2m-1} \sigma, \dots, x_m = \varrho^m \sigma^m, \\ x_{m+1} &= \varrho^{m-1} \sigma^{m+1}, \dots, x_{2m} = \sigma^{2m} \end{aligned}$$

hat. Es ist zu zeigen, daß auch umgekehrt jede Lösung des Systems (a), (b), (c) die Gestalt (6) hat. Nun hat jede Lösung von (a) nach Voraussetzung die Gestalt

$$(7) \quad x_0 = \varrho^m, \quad x_1 = \varrho^{m-1} \sigma, \dots, x_m = \sigma^m,$$

und eine Lösung von (b), bei der $x_m = \sigma^m$ ist, hat wegen der gleichen Voraussetzung das Aussehen

$$(8) \quad x_m = \sigma^m, \quad x_{m+1} = \sigma^{m-1} \tau, \dots, x_{2m} = \tau^m.$$

Nun soll aber auch die Gleichung (c) gelten, und wenn man in diese die Werte (7) und (8) einsetzt, ergibt sich

$$\varrho^m \tau^m - \binom{m}{1} \varrho^{m-1} \tau^{m-1} \sigma^2 + \binom{m}{2} \varrho^{m-2} \tau^{m-2} \sigma^4 - \dots + \binom{m}{m} \sigma^{2m} = 0,$$

oder also

$$(\varrho \tau - \sigma^2)^m = 0.$$

Daher ist $\sigma = \sqrt{\varrho} \sqrt{\tau}$, so daß die Lösung (7), (8) die folgende ist:

$$\begin{aligned} x_0 &= \sqrt{\varrho}^{2m}, \quad x_1 = \sqrt{\varrho}^{2m-1} \sqrt{\tau}, \dots, x_m = \sqrt{\varrho}^m \sqrt{\tau}^m, \\ x_{m+1} &= \sqrt{\varrho}^{m-1} \sqrt{\tau}^{m+1}, \dots, x_{2m} = \sqrt{\tau}^{2m}. \end{aligned}$$

Das geht aber durch die Bezeichnungsänderung $\sqrt{\varrho} \rightarrow \varrho, \sqrt{\tau} \rightarrow \sigma$ über in (6).

Jetzt nehmen wir umgekehrt bei einem gewissen Wert n an, die Kurve (1) lasse sich als Durchschnitt von $n - 1$ Flächen zweiter Ordnung darstellen:

$$f_1(x_0, x_1, \dots, x_n) = 0, \dots, f_{n-1}(x_0, x_1, \dots, x_n) = 0,$$

und wollen zeigen, daß dann n eine Potenz von 2 ist.

Zu dem Zweck seien u_0, u_1, \dots, u_n Unbestimmte und die Wurzeln der Gleichung

$$u_0 \varrho^n + u_1 \varrho^{n-1} + \dots + u_n = 0$$

seien $\varrho_1, \varrho_2, \dots, \varrho_n$. Dann hat das System der n Gleichungen

$$f_1 = 0, \dots, f_{n-1} = 0, \quad \sum u_i x_i = 0,$$

da die Lösung der ersten $n - 1$ Gleichungen nach Voraussetzung durch die Formeln (1) gegeben ist, die n Lösungen

$$(9) \quad x_0 = \varrho_i^n, \quad x_1 = \varrho_i^{n-1}, \dots, x_{n-1} = \varrho_i, \quad x_n = 1$$

für $i = 1, 2, \dots, n$ und keine anderen (abgesehen von einem Proportionalitätsfaktor). Daher ist, wenn v_0, v_1, \dots, v_n weitere Unbestimmte sind, die Resultante

$$R = R \left(\begin{matrix} f_1, \dots, f_{n-1}, \sum u_i x_i, \sum v_i x_i \\ x_0, \dots, x_{n-2}, x_{n-1}, x_n \end{matrix} \right)$$

nicht gleich 0, sondern ist in bezug auf die v_i ein homogenes Polynom vom Grad

$$2 \cdot 2 \dots 2 \cdot 1 = 2^{n-1}$$

und zerfällt in Linearfaktoren³⁾

$$(10) \quad R = C \prod_{i=1}^n (v_0 \varrho_i^n + v_1 \varrho_i^{n-1} + \dots + v_{n-1} \varrho_i + v_n)^{k_i},$$

wo k_i die Multiplizität der Lösung (9) ist und

$$(11) \quad k_1 + k_2 + \dots + k_n = 2^{n-1}.$$

C ist dabei ein Polynom der u_i . Nun ist aber die Resultante in bezug auf die u_i ebenfalls ein Polynom, also ein symmetrisches Polynom der Wurzeln $\varrho_1, \varrho_2, \dots, \varrho_n$. Daher bleibt die Formel (10) richtig, wenn man die ϱ_i beliebig permutiert, und da eine Zerlegung in Linearfaktoren nur auf eine Weise möglich ist, ergibt sich, daß alle Exponenten k_i einander gleich sind, etwa gleich k . Aus (11) folgt dann aber

$$nk = 2^{n-1}.$$

Daher ist n ein Teiler von 2^{n-1} , also eine Potenz von 2. W. z. b. w.

³⁾ Man sehe etwa des Verfassers Buch: Algebra, Band I, zweite Auflage, 1932, S. 290; oder: van der Waerden, Moderne Algebra II (1931), § 79; zweite Auflage, 1940; § 83.

Beweis des Reziprozitätsgesetzes für quadratische Reste.

Von

Karl Dörge in Köln.

Der folgende Beweis ist eng verwandt mit dem Frobeniusschen Beweis aus den Sitzungsber. d. Berl. Akad. 1914. Die Frobeniussche Darstellung hat den großen Vorzug der Symmetrie, die folgende Fassung ist hingegen etwas anschaulicher.

Es seien p, q zwei positive ungerade Primzahlen und $p > q$. v sei eine ganze Zahl. Unter B'_v verstehe man das — etwa abgeschlossene — Intervall der Länge $\frac{q}{2}$, dessen linker Endpunkt $v \cdot p$ ist, also $\left[v \cdot p, v \cdot p + \frac{q}{2} \right]$, und unter B''_v das — etwa abgeschlossene — Intervall von der Länge $\frac{p}{2}$, dessen rechter Endpunkt $v \cdot p$ ist, also $\left[v \cdot p - \frac{p}{2}, v \cdot p \right]$.



τ' sei die Anzahl derjenigen unter den Intervallen $B'_1, B'_2, \dots, B'_{\frac{p-1}{2}}$, in welchen ein Vielfaches von q liegt. Dann ist also τ' die Anzahl der Vielfachen von q in der Vereinigungsmenge der $B'_1, B'_2, \dots, B'_{\frac{p-1}{2}}$. Nach dem Gaußschen

Kriterium ist dann $\left(\frac{p}{q} \right) = (-1)^{\tau'}$. τ'' sei die Anzahl derjenigen unter den Vielfachen

$$1 \cdot q, 2 \cdot q, \dots, \frac{p-1}{2} \cdot q,$$

welche in die Vereinigungsmenge aller B'' hineinfallen. Dann ist nach dem Gaußschen Kriterium, wie man sofort sieht, $\left(\frac{q}{p} \right) = (-1)^{\tau''}$.

Ein B'_v mit $v \leq 0$ enthält nun sicher keine Zahl aus $*$. Der Endpunkt von $B''_{\frac{q-1}{2}}$ ist $p \cdot \frac{q-1}{2}$, der Anfangspunkt von $B''_{\frac{q-1}{2}+1}$ ist $p \cdot \frac{q-1}{2} + \frac{p}{2}$.

Wegen $p > q$ liegt zwischen ihnen die größte Zahl aus $*$: $\frac{p-1}{2} \cdot q$. Also ist τ'' die Anzahl aller Vielfachen von q in der Vereinigungsmenge der $B''_1, B''_2, \dots, B''_{\frac{q-1}{2}}$. Die Vereinigungsmenge

$$B'_1 + B'_2 + \dots + B'_{\frac{q-1}{2}} + B''_1 + B''_2 + \dots + B''_{\frac{q-1}{2}}$$

nenne man B . Jedes B' hat mit genau einem B'' einen Punkt gemeinsam, dieser ist aber ein Vielfaches von p , also gewiß nicht Zahl aus $*$. Unter τ verstehe man die Anzahl der Vielfachen von q in B . Dann ist also $\tau = \tau' + \tau''$.

Daraus folgt: Die Aussage $\left(\frac{p}{q}\right) \neq \left(\frac{q}{p}\right)$ ist äquivalent der Aussage: Von τ' , τ'' ist eine Zahl gerade, die andere ungerade, also äquivalent der Aussage: τ ist ungerade.

Man gehe nun vom Endpunkt von $B''_{\frac{q-1}{2}}$ um $\frac{p}{2}$ weiter nach rechts, man kommt zu einem Punkt, den man S nenne. Die Abszisse von S ist dann: $p \cdot \frac{q-1}{2} + \frac{q}{2} + \frac{p}{2} = \frac{p+1}{2} \cdot q$, also Vielfaches von q . Die Verteilung der Punkte von B auf das Intervall $[OS]$ hat nach dieser Konstruktion von S , welches vom letzten Punkte von B genau so weit nach rechts entfernt ist wie O nach links vom ersten Punkte von B , die folgende Eigenschaft: Geht man einerseits von O nach rechts und andererseits von S nach links, aber beidemale um die gleiche Strecke, dann liegen entweder diese beiden Punkte beide im Bereich B oder beide nicht. Also liegen auf der linken Hälfte des Intervalls $[OS]$ genau ebenso viele solcher Vielfache von q , die in B hineinfallen, wie auf der rechten Hälfte, wenn man hierbei beide Male den Mittelpunkt von $[OS]$, falls er Vielfaches von q ist und in B fällt, nicht mitzählt.

Daraus folgt: Die Aussage: τ ist ungerade, ist äquivalent mit der gleichzeitigen Gültigkeit der zwei Aussagen:

1. Die Abszisse des Mittelpunktes M von $[OS]$ ist Vielfaches von q , und
2. M liegt in B .

M hat die Abszisse $m = \frac{p+1}{4} \cdot q$. Daher ist 1. äquivalent mit $p \equiv -1 \pmod{4}$, unabhängig von der arithmetischen Natur von q .

Wann gilt 2.? Ist $q = 4n + 1$, so ist $m = (4n + 1) \cdot \frac{p+1}{4} = np + \frac{p}{4} + \frac{q}{4}$, also liegt wegen $p > q$ hier M zwischen dem Endpunkt von B'_n und dem Anfangspunkt von B''_{n+1} , also nicht in B . Ist dagegen $q = 4n - 1$, so ist $m = (4n - 1) \cdot \frac{p+1}{4} = np - \frac{p}{4} + \frac{q}{4}$, also M in B'_n , also in B . Also ist 2. äquivalent mit $q \equiv -1 \pmod{4}$, unabhängig von der arithmetischen Natur von p .

Mithin folgt zusammengefaßt:

Die Aussage $\left(\frac{p}{q}\right) \neq \left(\frac{q}{p}\right)$ ist äquivalent mit der Doppelaussage

1. $p \equiv -1 \pmod{4}$, 2. $q \equiv -1 \pmod{4}$.

Über eine Metrik im Siegelschen Halbraum.

Von

Hans Maaß in Heidelberg.

Es sei Z ein n -reihige quadratische komplexe Matrix, $Z = X + iY$ die Zerlegung von Z in Real- und Imaginärteil. Für eine komplexe Matrix Ω soll $\Omega > 0$ bedeuten, daß Ω hermitisch und die zu Ω gehörige hermitische Form positiv definit ist. Der Siegelsche Halbraum \mathfrak{H} wird dann durch

$$(1) \quad Z' = Z, \quad Y > 0$$

beschrieben. Alle reellen Modulsstitutionen:

$$(2) \quad \sigma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \text{ mit } \sigma' \iota \sigma = \iota, \quad \iota = \begin{pmatrix} 0 & E \\ -E & 0 \end{pmatrix}$$

führen \mathfrak{H} vermöge

$$Z \rightarrow (AZ + B)(CZ + D)^{-1}$$

bekanntlich¹⁾ in sich über. In vorliegender Note wird eine Metrik für \mathfrak{H} eingeführt, welche als direkte Verallgemeinerung der bekannten hyperbolischen Maßbestimmung der Halbebene $\Im m z > 0$ anzusprechen ist. M. Sugawara hat sich bereits um eine solche Verallgemeinerung bemüht²⁾; er hat indessen den Ansatz, der hier zugrunde liegt, nicht weiter verfolgt. Wir gehen aus von der quadratischen Differentialform

$$(3) \quad ds^2 = Sp(dZY^{-1}\overline{dZ}Y^{-1}) \quad (Sp = \text{Spur});$$

sie definiert ein gegenüber allen reellen Modulsstitutionen invariantes Bogenelement ds für \mathfrak{H} . Für die durch ds induzierte Metrik werden im einzelnen die folgenden Sätze bewiesen. Unter allen (analytischen) Kurven, welche zwei gegebene Punkte von \mathfrak{H} innerhalb \mathfrak{H} verbinden, gibt es genau eine mit kürzester Länge. Diese ausgezeichnete Kurve (Geodätische) kann in folgendem Sinne explizit angegeben werden. Transformiert man, was stets möglich ist, zwei gegebene Punkte von \mathfrak{H} mittels einer geeignet gewählten

¹⁾ C. L. Siegel, Einführung in die Theorie der Modulfunktionen n -ten Grades. Math. Annalen 116 (1939), S. 617–657.

²⁾ M. Sugawara, On the General Zetafuchsian Functions. Proc. Acad. Tokyo 16 (1940), S. 367–372; A Generalization of Poincare-space. Proc. Acad. Tokyo 16 (1940), S. 373–377. Zur zweiten Arbeit ist zu bemerken, daß die Geodätischen bezüglich der besprochenen Metrik im allgemeinen keineswegs eindeutig bestimmt sind. Die Auszeichnung von „geraden Linien“ (straight lines) bedeutet daher eine gewisse Willkür.

reellen Modulsstitution in die spezielle Lage iE , iD (D = reelle Diagonalmatrix), so wird die Geodätische $Z(t) = (x_{\mu\nu}(t)) + i(y_{\mu\nu}(t))$, welche die Punkte iE , $iD = i(\delta_{\mu\nu}d_\nu)$ verbindet ($0 \leq t \leq 1$, $Z(0) = iE$, $Z(1) = iD$), durch das Gleichungssystem

$$(4) \quad x_{\mu\nu}(t) = 0, \quad y_{\mu\nu}(t) = \delta_{\mu\nu}d_\nu^t \quad \text{für } \mu, \nu = 1, 2, \dots, n, \quad 0 \leq t \leq 1$$

($\delta_{\mu\nu}$ = Kroneckersymbol)

beschrieben. Der Abstand zweier beliebiger Punkte Z_1, Z_2 aus \mathfrak{H} , d. h. die Länge der Geodätischen, welche Z_1 mit Z_2 verbindet, hat den Wert

$$(5) \quad s(Z_1, Z_2) = \sqrt{\sum_{\nu=1}^n \left(\log \frac{1+\lambda_\nu}{1-\lambda_\nu} \right)^2};$$

dabei sind λ_ν^2 , $\nu = 1, 2, \dots, n$ die charakteristischen Wurzeln von

$$(5a) \quad (Z_2 - Z_1)(Z_2 - \bar{Z}_1)^{-1}(\bar{Z}_2 - \bar{Z}_1)(\bar{Z}_2 - Z_1)^{-1}.$$

Für eine Reihe von Überlegungen ist es zweckmäßig, den Bereich \mathfrak{H} mittels

$$(6) \quad W = \tau(Z) = (Z - iE)(Z + iE)^{-1}, \quad \tau = \frac{1}{\sqrt{2}i} \begin{pmatrix} E & -iE \\ E & iE \end{pmatrix}$$

in den Bereich \mathfrak{E} .

$$(7) \quad W' = W, \quad E - W'\bar{W} > 0$$

zu transformieren³⁾. Entsprechend ist die Gruppe Γ der reellen Modulsstitutionen durch die transformierte Gruppe $\Gamma_1 = \tau\Gamma\tau^{-1}$ zu ersetzen. Die Substitutionen $\sigma \in \Gamma_1$ sind dadurch zu charakterisieren, daß

$$(8) \quad \sigma' i \sigma = i, \quad \sigma' x \bar{\sigma} = x \quad \text{mit } x = \begin{pmatrix} E & 0 \\ 0 & -E \end{pmatrix}.$$

Aus der Darstellung³⁾ entnimmt man zunächst leicht den folgenden Sachverhalt. W_1, W_2 seien zwei beliebige Punkte aus \mathfrak{E} , λ_ν^2 , $\nu = 1, 2, \dots, n$, die charakteristischen Wurzeln von

$$(9) \quad (W_2 - W_1)(E - \bar{W}_1 W_2)^{-1}(\bar{W}_2 - \bar{W}_1)(E - W_1 \bar{W}_2)^{-1},$$

dann ist jede symmetrische Funktion der Wurzeln λ_ν^2 , $\nu = 1, 2, \dots, n$, insbesondere

$$s(W_1, W_2) = \sqrt{\sum_{\nu=1}^n \left(\log \frac{1+\lambda_\nu}{1-\lambda_\nu} \right)^2}$$

³⁾ Vgl. M. Sugawara, Über eine allgemeine Theorie der Fuchsschen Gruppen und Theta-Reihen. *Annals of Math.* 41 (1940), S. 488–494.

invariant gegenüber den Substitutionen von Γ_1 . Zwischen den Punkten W_k und ihren Originalen $Z_k = \tau^{-1}(W_k)$, $k = 1, 2$, bestehen, wie man leicht ein-
sieht, die Beziehungen

$$\begin{aligned} W_2 - W_1 &= +2i(Z_1 + iE)^{-1}(Z_2 - Z_1)(Z_2 + iE)^{-1}, \\ E - \bar{W}_1 W_2 &= -2i(\bar{Z}_1 - iE)^{-1}(Z_2 - \bar{Z}_1)(Z_2 + iE)^{-1}, \\ \bar{W}_2 - \bar{W}_1 &= -2i(\bar{Z}_1 - iE)^{-1}(\bar{Z}_2 - \bar{Z}_1)(\bar{Z}_2 - iE)^{-1}, \\ E - W_1 \bar{W}_2 &= +2i(Z_1 + iE)^{-1}(\bar{Z}_2 - Z_1)(\bar{Z}_2 - iE)^{-1}. \end{aligned} \quad (10)$$

Verknüpfen wir diese vier Gleichungen derart, daß linksseitig die Matrix (9) entsteht, so erhalten wir rechtsseitig die mit $(Z_1 + iE)$ transformierte Matrix (5a). Die beiden Matrizen (5a) und (9) haben daher dieselben charakteristischen Wurzeln, woraus die Invarianz von (5) bezüglich Γ erhellt. Für zwei infinitesimal benachbarte Punkte Z , $Z + dZ$ liefert (5) unser Bogenelement:

$$s(Z, Z + dZ) = \sqrt{\sum_{\nu=1}^n (2\lambda_\nu)^2} = \sqrt{Sp(dZ Y^{-1} d\bar{Z} Y^{-1})} = ds.$$

Dieses ist also ebenfalls gegenüber den reellen Modulsstitutionen invariant. Wir brauchen nur noch einzusehen, daß $s(Z_1, Z_2)$ für die durch (3) definierte Metrik die oben angegebene Bedeutung hat. Das ergibt sich automatisch am Schluß der Betrachtung. Zunächst wollen wir uns davon überzeugen, daß zwei gegebene Punkte von \mathfrak{H} durch eine geeignet gewählte Substitution aus Γ stets in die oben angegebene spezielle Lage gebracht werden können. Es genügt natürlich, die entsprechende Aufgabe im transformierten Bereich \mathfrak{G} zu lösen: Zwei gegebene Punkte W_1, W_2 sind durch eine geeignet gewählte Substitution aus Γ_1 in die Lage O, D_1 (= reelle Diagonalmatrix) zu transformieren. Nach ³⁾ ist es möglich, den Punkt W_1 in den Nullpunkt O zu transformieren. Demnach können wir uns offenbar auf den Fall $W_1 = O$ beschränken. Gesucht wird dann eine Substitution $\sigma \in \Gamma_1$, für welche

$$\sigma(O) = O, \quad \sigma(W_2) = D_1.$$

Die Substitution σ hat notwendig die Gestalt

$$\sigma = \begin{pmatrix} U & O \\ O & \bar{U} \end{pmatrix} \quad \text{mit} \quad U' \bar{U} = E.$$

Es muß also eine unitäre Matrix U bestimmt werden, für welche $\sigma(W_2) = UW_2 \bar{U}^{-1} = UW_2 U' = D_1$ eine reelle Diagonalmatrix wird. Eine solche Matrix U gibt es in der Tat für jede symmetrische komplexe Matrix W_2 ⁴⁾. Wir kommen nun zum Eindeutigkeitsbeweis für die Geodätische durch zwei

⁴⁾ U. Wegner u. J. Wellstein, Bemerkungen zur Transformation von komplexen symmetrischen Matrizen. Monatshefte f. Math. u. Phys. 40 (1933), S. 319–322.

beliebige Punkte Z_1, Z_2 aus \mathfrak{H} . Ohne Beschränkung der Allgemeinheit können und wollen wir im folgenden $Z_1 = iE, Z_2 = iD, D =$ reelle Diagonalmatrix mit den Diagonalelementen $d_\nu, \nu = 1, 2, \dots, n$, voraussetzen. $Z = Z(t) = (x_\nu(t)) + i(y_\nu(t)), 0 \leq t \leq 1$, sei die Gleichung für die vorerst noch hypothetische Geodätische zwischen Z_1 und Z_2 und

$$(11) \quad s(Z_1, Z_2) = \int_0^1 \sqrt{Sp(\dot{Z} Y^{-1} \bar{\dot{Z}} Y^{-1})} dt \quad (\dot{() = \frac{d}{dt}()})$$

ihre Länge. Da man bei der Spurbildung eines Matrizenproduktes die Faktoren zyklisch vertauschen kann, so zerfällt die Form (3) in

$$(12) \quad Sp(\dot{Z} Y^{-1} \bar{\dot{Z}} Y^{-1}) = Sp(\dot{X} Y^{-1})^2 + Sp(\dot{Y} Y^{-1})^2.$$

Die beiden Summanden auf der rechten Seite dieser Gleichung sind positiv definite quadratische Formen in den Elementen von \dot{X} bzw. \dot{Y} . Es muß nun notwendig $Sp(\dot{X} Y^{-1})^2 = 0$ und daher $\dot{X} = 0$ sowie $X = 0$ identisch in t sein, da man anderenfalls in $iY(t)$ eine kürzere Verbindung von Z_1 und Z_2 an Stelle von $Z(t)$ hätte, was der Voraussetzung über $Z(t)$ widerspricht. Der nächste Schritt soll zeigen, daß $Y(t)$ für jeden Wert von t Diagonalgestalt haben muß. Wir bestimmen zu diesem Zweck eine reelle orthogonale Matrix $U = U(t)$ derart, daß $H = H(t) = U^{-1} Y U'^{-1}$ eine Diagonalmatrix mit den Diagonalelementen h_1, h_2, \dots, h_n wird. Dieses Verfahren kann so durchgeführt werden, daß $U(0) = E, H(1) = D$ und die Elemente von U und H analytische Funktionen von t werden, sofern nur vorausgesetzt wird, daß $Z(t)$ von t analytisch abhängt, was hiermit geschehen soll. Mit Hilfe von $U U' = E$ findet man:

$$\begin{aligned} \dot{Y} Y^{-1} &= \dot{U} U^{-1} + U \dot{H} H^{-1} U^{-1} + U H \dot{U}' U H^{-1} U^{-1}, \\ Sp(\dot{Y} Y^{-1})^2 &= Sp(\dot{U} U^{-1})^2 + Sp(\dot{H} H^{-1})^2 + Sp(\dot{U}' U)^2 + \\ &+ 2 Sp(\dot{U} \dot{H} H^{-1} U^{-1}) + 2 Sp(\dot{H} \dot{U}' U H^{-1}) + 2 Sp(\dot{U} \dot{H} \dot{U}' U H^{-1} U^{-1}). \end{aligned}$$

Zur Abkürzung wird $\Omega = (\omega_\mu) = U^{-1} \dot{U}$ angeführt. $\dot{U} U' + U \dot{U}' = 0$ besagt, daß $\Omega' = -\Omega$. Die Schiefsymmetrie von Ω hat insbesondere

$$\begin{aligned} Sp(\dot{U} \dot{H} H^{-1} U^{-1}) &= Sp(\Omega \dot{H} H^{-1}) = 0, \\ Sp(\dot{H} \dot{U}' U H^{-1}) &= -Sp(\dot{H} \Omega H^{-1}) = 0 \end{aligned}$$

zur Folge und damit schließlich die einfache Gleichung

$$\begin{aligned} Sp(\dot{Y} Y^{-1})^2 &= -2 Sp(\Omega \Omega') + 2 Sp(\Omega H \Omega' H^{-1}) + Sp(\dot{H} H^{-1})^2 \\ (13) \quad &= 2 \sum_{\mu, \nu=1}^n \omega_\mu^2 (h_\mu h_\nu^{-1} - 1) + Sp(\dot{H} H^{-1})^2 \\ &= 2 \sum_{\mu < \nu} \omega_\mu^2 (h_\mu h_\nu^{-1} + h_\nu h_\mu^{-1} - 2) + Sp(\dot{H} H^{-1})^2. \end{aligned}$$

Das Minimum von $ab^{-1} + ba^{-1}$ für $a, b > 0$ liegt bei 2 und wird nur für $a = b$ angenommen. Wir stellen damit fest: Es ist stets

$$(14) \quad Sp(\dot{Y}Y^{-1})^2 \geq Sp(\dot{H}H^{-1})^2,$$

und das Gleichheitszeichen gilt genau dann, wenn aus $h_\nu \neq h_\mu$ stets $\omega_{\mu\nu} = 0$ folgt. Das Gleichheitszeichen muß aber gelten, da andernfalls die Verbindungskurve $iH(t)$ von Z_1 nach Z_2 wieder kleinere Länge als die hypothetische Geodätische $Z(t)$ hätte. Unter den charakteristischen Wurzeln $\lambda_1, \lambda_2, \dots, \lambda_n$ von $Y(t)$ mögen im allgemeinen p untereinander verschiedene vorkommen. Wir bezeichnen sie mit $\lambda_\mu = \lambda_\mu(t)$, $\mu = 1, 2, \dots, p$; n_μ sei die Vielfachheit der allgemeinen Wurzel λ_μ . Wegen der Analytizität kann $h_\mu(t) = \lambda_\mu(t)$ für $\mu \neq \nu$ nur in endlich vielen Punkten t stattfinden. Wir dürfen uns die Wurzeln h_1, h_2, \dots, h_n so angeordnet denken, daß die n_1 ersten Wurzeln dieser Reihe gleich λ_1 , die n_2 folgenden Wurzeln gleich λ_2 usw., die n_p letzten Wurzeln gleich λ_p sind. Im allgemeinen hat dann Ω die Gestalt

$$(15) \quad \Omega = \begin{pmatrix} \Omega_1 & & 0 \\ & \Omega_2 & \\ 0 & & \ddots & \\ & & & \Omega_p \end{pmatrix},$$

wobei Ω_μ eine quadratische Matrix von n_μ Zeilen ($\mu = 1, 2, \dots, p$). Da Ω analytisch von t abhängt, gilt Gleichung (15) für alle t . Für die orthogonale Matrix U gilt entsprechend zu (15)

$$(16) \quad U = \begin{pmatrix} U_1 & & 0 \\ & U_2 & \\ 0 & & \ddots & \\ & & & U_p \end{pmatrix};$$

denn bei gegebener Matrix Ω ist die Matrixdifferentialgleichung

$$\dot{U} = U\Omega$$

mit der Anfangsbedingung $U(0) = E$ eindeutig auflösbar, und es gibt eine Lösung der Gestalt (16). Selbstverständlich sind alle Matrizen U_μ ebenfalls orthogonal, so daß also

$$(17) \quad Y(t) = H(t).$$

Die Gleichung (11) vereinfacht sich damit zu

$$(18) \quad s(Z_1, Z_2) = \int_0^1 \sqrt{\sum_{r=1}^n \dot{y}_{r,r}^2} dt.$$

Deutet man $(\log y_{11}, \log y_{22}, \dots, \log y_{nn})$ als cartesische Koordinaten eines Punktes im n -dimensionalen Raum R_n , so wird $s(Z_1, Z_2)$ gleich der euklidischen

Länge der Integrationskurve. Das absolute Minimum von $s(Z_1, Z_2)$ erhält man also nur dann, wenn man über die Gerade im R_n integriert, welche die Punkte mit den Koordinaten $(0, 0, \dots, 0)$ und $(\log d_1, \log d_2, \dots, \log d_n)$ miteinander verbindet. Der Eindeutigkeitsbeweis für die Geodätische ist damit erbracht, die Existenz natürlich auch gesichert. Überdies erhalten wir die noch fehlenden Geodätengleichungen

$$\log y_v = t \log d_v \quad \text{für } v = 1, 2, \dots, n, \quad 0 \leq t \leq 1,$$

und die Länge der Geodätischen

$$(19) \quad s(Z_1, Z_2) = \sqrt{\sum_{v=1}^n (\log d_v)^2}.$$

Die charakteristischen Wurzeln λ_v^2 von (5a) sind im vorliegenden Fall gleich $\left(\frac{d_v-1}{d_v+1}\right)^2$; damit ergibt sich die Übereinstimmung von (5) und (19). Wegen der bereits erkannten Invarianz von (5) gegenüber den reellen Modulusubstitutionen liefert Formel (5) den Abstand der Punkte Z_1, Z_2 auch bei allgemeiner Lage der Punkte.

Bettet man die Mannigfaltigkeit \mathfrak{H} in den $n(n+1)$ -dimensionalen reellen Raum der Punkte mit den cartesischen Koordinaten $x_{\mu\nu}, y_{\mu\nu}, \mu \leq \nu, \mu, \nu = 1, 2, \dots, n$ ein, so erhält man eine euklidische Metrik für \mathfrak{H} gemäß der Abstandsformel

$$(20) \quad e(Z^*, Z) = \sqrt{\sum_{\mu \leq \nu} (z_{\mu\nu}^* - z_{\mu\nu})(\overline{z_{\mu\nu}^* - z_{\mu\nu}})}.$$

Für gewisse Überlegungen ist es von Interesse zu wissen, in welcher Beziehung die Umgebungssysteme der beiden Maßbestimmungen zueinander stehen. Im Falle $n = 1$ sind die Verhältnisse vollständig zu übersehen. Es ist z. B. leicht einzusehen, daß für einen Punkt $z = x + iy$ der oberen z -Halbebene die Ungleichungen

$$(21) \quad \begin{aligned} s(z, z + \Delta z) &\leq \frac{1}{1-\delta} \left| \frac{\Delta z}{y} \right| \quad \text{für} \quad \left| \frac{\Delta z}{y} \right| \leq \delta < 1, \\ \left| \frac{\Delta z}{y} \right| &\leq \frac{2}{2-3\delta} s(z, z + \Delta z) \quad \text{für} \quad s(z, z + \Delta z) \leq \delta < \frac{2}{3} \end{aligned}$$

bestehen. Diese Ungleichungen haben ihre Analoga für allgemeines n . Für $Z = X + iY, Y = C^*C, Z = C^*Z^*C, \Delta Z = C^*\Delta Z^*C, 0 < \varepsilon \leq \varepsilon_0(n)$ und gewisse nur von n abhängige Konstanten c_1, c_2 gilt nämlich

$$(22) \quad \begin{aligned} s(Z, Z + \Delta Z) &\leq \varepsilon \quad \text{für} \quad e(Z^*, Z^* + \Delta Z^*) \leq c_1 \varepsilon, \\ e(Z^*, Z^* + \Delta Z^*) &\leq \varepsilon \quad \text{für} \quad s(Z, Z + \Delta Z) \leq c_2 \varepsilon. \end{aligned}$$

Diese Abschätzungen sind aus (5) abzuleiten:

$$\begin{aligned}
 s(Z, Z + \Delta Z) &= \sqrt{\sum_{v=1}^n \left(\log \frac{1+\lambda_v}{1-\lambda_v} \right)^2}, \\
 (23) \quad \sum_{v=1}^n \lambda_v^2 &= Sp \Delta Z^* (\Delta Z^* + 2iE)^{-1} \overline{\Delta Z^*} (\overline{\Delta Z^*} - 2iE)^{-1} \\
 &= Sp \Xi \overline{\Xi} = \sum_{\mu, \nu=1}^n \xi_{\mu\nu} \overline{\xi_{\mu\nu}}
 \end{aligned}$$

mit

$$(24) \quad \Xi = \Xi' = \Delta Z^* (\Delta Z^* + 2iE)^{-1} = (\xi_{\mu\nu}).$$

Die letzte Gleichung ergibt nach ΔZ^* aufgelöst:

$$(25) \quad \Delta Z^* = 2i\Xi(E - \Xi)^{-1} = (\Delta z_{\mu\nu}^*).$$

Mit $O(\varepsilon)$ bezeichnen wir allgemein einen Ausdruck, welcher nach Division durch ε dem Betrage nach unter einer nur von n abhängigen Schranke liegt, sofern man sich auf ein hinreichend kleines ε -Intervall $0 < \varepsilon \leq \varepsilon_0(n)$ beschränkt. Nun zum Beweis von (22):

1. Sei $e(Z^*, Z^* + \Delta Z^*) \leq \varepsilon$, dann folgt der Reihe nach
 $\Delta z_{\mu\nu}^* = O(\varepsilon)$, $\xi_{\mu\nu} = O(\varepsilon)$, $\lambda_v = O(\varepsilon)$, $s(Z, Z + \Delta Z) = O(\varepsilon)$.
2. Sei $s(Z, Z + \Delta Z) \leq \varepsilon$, dann ergibt sich nacheinander
 $\frac{1+|\lambda_v|}{1-|\lambda_v|} \leq e^{\varepsilon}$, $|\lambda_v| \leq \frac{e^{\varepsilon}-1}{e^{\varepsilon}+1} = O(\varepsilon)$, $\xi_{\mu\nu} = O(\varepsilon)$, $\Delta z_{\mu\nu}^* = O(\varepsilon)$,
 $e(Z^*, Z^* + \Delta Z^*) = O(\varepsilon)$.

(Eingegangen am 23. 9. 1941.)

Simultane Darstellung zweier ganzen Zahlen als einer Summe von ganzen Zahlen und deren Quadratsumme.

Von

H. D. Kloosterman in Leiden.

In dieser Arbeit wird ein Anfang gemacht mit dem Studium des diophantischen Gleichungssystems

$$(1) \quad Q(x) = \sum_{i,j=1}^s a_{ij} x_i x_j = n, \quad L(x) = \sum_{i=1}^s b_i x_i = m.$$

Hier sind a_{ij} ($i, j = 1, 2, \dots, s$), b_i ($i = 1, 2, \dots, s$), n und m gegebene ganze Zahlen. Hauptsächlich interessiert uns dabei die Anzahl der Systeme von ganzen Zahlen x_1, x_2, \dots, x_s , welche dem Gleichungssystem (1) genügen. Diese Anzahl werde mit $r(n, m)$ bezeichnet.

Es seien $w = e^{\pi i \tau}$ und v komplexe Veränderliche, und man betrachte die Funktion

$$(2) \quad f(v|\tau) = \sum_{x_1, x_2, \dots, x_s = -\infty}^{+\infty} w^{Q(x)} e^{2\pi i v L(x)}$$

der beiden Veränderlichen v und τ . Falls die quadratische Form $Q(x)$ positiv-definit ist (was wir weiterhin immer annehmen werden), so ist die s -fache unendliche Reihe (2) für $\Im(\tau) > 0$ absolut konvergent, und es ist deshalb auch:

$$f(v|\tau) = \sum_{m, n = -\infty}^{+\infty} r(n, m) w^n e^{2\pi i v m}.$$

Demnach ist

$$(3) \quad r(n, m) = \frac{1}{2\pi i} \int \frac{dw}{w^{n+1}} \int_0^1 f(v|\tau) e^{-2\pi i v m} dv,$$

wo das erste Integral über einen Kreis Γ mit Radius $R < 1$ um den Nullpunkt der w -Ebene zu erstrecken ist. Es ist nun leicht, mit Hardy-Littlewood'schen Methoden aus der Formel (3) eine Näherungsformel für $r(n, m)$ zu gewinnen. Dies ist indessen nicht der Zweck der vorliegenden Arbeit. Vielmehr soll gezeigt werden, daß der für $r(n, m)$ erhaltene Näherungsausdruck in einigen Spezialfällen zu *exakten* Formeln für $r(n, m)$ führt. Dies wird gezeigt mit Methoden, die auf Mordell und Hardy zurückgehen¹⁾. Der behandelte Spezialfall beschäftigt sich mit dem Gleichungssystem

$$(4) \quad x_1^2 + x_2^2 + \dots + x_s^2 = n, \quad x_1 + x_2 + \dots + x_s = m,$$

¹⁾ L. J. Mordell, On the representation of numbers as a sum of $2r$ squares. Quarterly Journ. of Math. 48 (1917), S. 93–104; G. H. Hardy, On the representation of a number as the sum of any number of squares, and in particular of five. Transact. Amer. math. Soc. 21 (1920), S. 255–284.

wobei die Fälle $s = 3, 5$ und 7 näher betrachtet werden sollen. In diesen Fällen ergeben sich exakte Formeln für $r(n, m)$. Falls man in den erhaltenen Formeln noch m veränderlich nimmt, erhält man neue Formeln für die Anzahl der Darstellungen natürlicher Zahlen als eine Summe von $3, 5$ oder 7 Quadraten.

Zusatz bei der Korrektur. Durch Elimination einer Variablen aus den beiden Gleichungen (4) hätte man das Problem auch zurückführen können auf das bekannte Problem der Darstellung natürlicher Zahlen durch eine gewisse quadratische Form in $s - 1$ Variablen, die dann noch gewissen Kongruenzbedingungen genügen müssen. Oder analytisch gesprochen: man hätte auch Funktionen einer Veränderlichen (nämlich Modulformen der Stufe s) statt den Funktionen (2) in zwei Veränderlichen zur Lösung des Problems heranziehen können. Dies hätte den Vorteil gehabt, daß man die ganze Theorie der Modulformen, wie sie heute vorliegt, hätte benutzen können. In dieser Arbeit ist aber der direkte Weg gewählt, indem Funktionen in zwei Veränderlichen benutzt werden, welche einer Funktionalgleichung genügen von der Art, wie sie für $\theta(v|\tau)$ gilt. Funktionentheoretisch erhält man dann Identitäten, wie z. B. aus den Sätzen 3, 10 und 11 dieser Arbeit einige hervorgehen. Wie aus dem obigen hervorgeht, hätte man diese Identitäten auch über den „arithmetischen“ Umweg erhalten können, der mit der oben erwähnten Elimination einer Variablen aus den beiden Gleichungen (4) gleichwertig ist. Es erscheint aber wünschenswert, direkte funktionentheoretische Beweise zu finden und es entsteht das Problem, die speziellen oben erwähnten Identitäten in eine allgemeinere Theorie einzuordnen. Verf. hofft in einer späteren Arbeit auf dieses Problem zurückzukommen. In der vorliegenden Arbeit sind vorläufig noch die älteren, von Mordell und Hardy stammenden Beweismethoden benutzt.

Bezeichnungen. Es bedeuten

$$\sum_{\lambda \bmod q} \quad \text{bzw.} \quad \sum'_{\lambda \bmod q}$$

Summen, in denen die Summationsveränderliche λ ein vollständiges bzw. ein teilerfremdes Restsystem mod q durchläuft.

Zur Abkürzung wird oft $e(x)$ statt $e^{2\pi i x}$ geschrieben. Es ist also $e(x)$ eine Funktion mit der Periode 1.

$a|b$ bedeutet, daß a in b aufgeht (a und b ganz).

Der größte gemeinsame Teiler zweier ganzen Zahlen a und b wird mit (a, b) bezeichnet. Ebenso wird der größte gemeinsame Teiler dreier ganzen Zahlen a, b, c mit (a, b, c) bezeichnet. Der größte gemeinsame Teiler ist immer als natürliche Zahl gemeint.

Kongruenzen $a \equiv b \pmod{n}$ werden oft (besonders, wenn sie als Summationsbedingungen unter einem Σ -Zeichen geschrieben werden) zu $a \equiv b \pmod{n}$ abgekürzt.

$\left(\frac{a}{b}\right)$ ist (für ungerades $b \geq 1$ und ganzes a) das Symbol von Jacobi aus der Theorie der quadratischen Reste. Falls $(a, b) > 1$ ist, wird das Symbol $\left(\frac{a}{b}\right)$ als 0 definiert. Für jede ganze Zahl a (auch für $a = 0$) ist $\left(\frac{a}{1}\right) = 1$ zu nehmen.

$\mu(n)$ ist die zahlentheoretische Funktion von Moebius.

$\operatorname{sgn} x$ bedeutet $\frac{x}{|x|}$ für $x \neq 0$ und 0 für $x = 0$.

Eine zahlentheoretische Funktion $f(n)$ wird *multiplikativ* genannt, falls für jedes Paar teilerfremder ganzen Zahlen n_1 und n_2 gilt:

$$f(n_1 n_2) = f(n_1) f(n_2).$$

§ 1.

Formale Konstruktion der singulären Reihe.

Weiterhin sei immer $r(n, m) = r_s(n, m)$ die Lösungszahl des diophantischen Gleichungssystems (4). In diesem Spezialfall reduziert sich die Funktion (2) auf

$$f(v|\tau) = \vartheta^s(v|\tau),$$

wo

$$(5) \quad \vartheta(v|\tau) = \sum_{x=-\infty}^{+\infty} w^x e^{2\pi i \tau x}$$

ist. Aus (3) wird demnach

$$(6) \quad r(n, m) = \frac{1}{2\pi i} \int \frac{dw}{w^{n+1}} \int_0^1 \vartheta^s(v|\tau) e^{-2\pi i v m} dv.$$

Es werde nun der Kreis Γ in Farey-Bogen einer gewissen Ordnung N zerteilt. Diese Farey-Bogen sind den rationalen Brüchen $\frac{a}{q}$ mit ganzen a und q und

$$0 < a \leq q \leq N, \quad (a, q) = 1$$

eindeutig zugeordnet. Der Farey-Bogen $\xi_{a,q}$, der dem Bruch $\frac{a}{q}$ zugeordnet ist, wird durch

$$w = e^{\frac{2\pi i a}{q}} - \frac{1}{n} + i\theta, \quad -\frac{2\pi}{q(q+q'')} \leq \theta \leq \frac{2\pi}{q(q+q')}$$

definiert, falls $\frac{a'}{q'}$ bzw. $\frac{a''}{q''}$ der nächstgrößere bzw. nächstkleinere Bruch zu $\frac{a}{q}$

aus der Farey-Reihe der Ordnung N ist. Es ist also $R = e^{-\frac{1}{n}}$ gewählt. Falls $a = q = 1$ ist, ist dabei $\xi_{a,q}$ als die Summe der beiden durch

$$w = e^{-\frac{1}{n} + i\theta}, \quad 0 \leq \theta \leq \frac{2\pi}{N+1}, \quad 2\pi - \frac{2\pi}{N+1} \leq \theta \leq 2\pi$$

definierten Bogen aufzufassen.

Das Segment $0 \leq v \leq 1$ wird ebenfalls in Farey-Segmente einer gewissen Ordnung M zerteilt, welche den rationalen Brüchen $\frac{b}{r}$ mit ganzen b und r und

$$0 < b \leq r \leq M, \quad (b, r) = 1$$

eindeutig zugeordnet sind. Das Farey-Segment $\eta_{b,r}$, das dem Bruch $\frac{b}{r}$ zugeordnet ist, wird durch

$$v = \frac{b}{r} + \beta, \quad -\frac{1}{r(r+r'')} \leq \beta \leq \frac{1}{r(r+r')}$$

definiert, falls $\frac{b'}{r'}$ bzw. $\frac{b''}{r''}$ der nächstgrößere bzw. nächstkleinere Bruch zu $\frac{b}{r}$ aus der Farey-Reihe der Ordnung M ist. Falls $b = r = 1$ ist, ist dabei $\eta_{b,r}$ als die Summe der beiden durch

$$0 \leq v \leq \frac{1}{M+1} \quad \text{und} \quad \frac{M}{M+1} \leq v \leq 1$$

definierten Segmente aufzufassen.

Auf gegebenen $\xi_{a,q}$ und $\eta_{b,r}$ wird $\vartheta(v|\tau)$ nun folgendermaßen approximiert. Schreibt man

$$(7) \quad w = e\left(\frac{a}{q}\right) W, \quad W = e^{-\frac{1}{n} + i\theta},$$

so wird

$$\vartheta(v|\tau) = \sum_{x=-\infty}^{+\infty} e\left(\frac{ax^2}{q} + \frac{bx}{r}\right) Wx^2 e^{2\pi i x \beta}.$$

Schreibt man jetzt

$$x = kqr + l,$$

wo k sämtliche ganze Zahlen und l die Zahlen $0, 1, 2, \dots, qr - 1$ durchläuft so wird

$$\vartheta(v|\tau) = \sum_{l=0}^{qr-1} e\left(\frac{al^2}{q} + \frac{bl}{r}\right) \sum_{k=-\infty}^{+\infty} W(kqr+l)^2 e^{2\pi i \beta(kqr+l)}$$

und also unter Benutzung von (7):

$$(8) \quad \vartheta(v|\tau) = \sum_{l=0}^{qr-1} e\left(\frac{al^2}{q} + \frac{bl}{r}\right) W^2 e^{2\pi i \beta l} \times \\ \times \vartheta\left(\frac{qrl\left(-\frac{1}{n} + i\theta\right)}{\pi i} + qr\beta \left| \frac{q^2 r^2 \left(-\frac{1}{n} + i\theta\right)}{\pi i} \right.\right).$$

Es ist aber

$$\vartheta(v|\tau) = \frac{1}{\sqrt{-i\tau}} e^{-\frac{\pi i v^2}{\tau}} \vartheta\left(\frac{v}{\tau} \middle| -\frac{1}{\tau}\right) \quad (\Re \sqrt{-i\tau} > 0).$$

Ersetzt man hier

$$v \text{ durch } \frac{qr l \left(-\frac{1}{n} + i\Theta\right)}{\pi i} + qr\beta \quad \text{und} \quad \tau \text{ durch } \frac{q^2 r^2 \left(-\frac{1}{n} + i\Theta\right)}{\pi i},$$

so wird (8):

$$\begin{aligned} \vartheta(v|\tau) = \frac{1}{qr} \sqrt{\frac{\pi}{\frac{1}{n} - i\Theta}} e^{-\frac{\pi^2 \beta^2}{\frac{1}{n} - i\Theta}} \sum_{l=0}^{qr-1} e\left(\frac{al^2}{q} + \frac{bl}{r}\right) \times \\ \times \vartheta\left(\frac{l}{qr} + \frac{\beta \pi i}{qr\left(-\frac{1}{n} + i\Theta\right)} \middle| \frac{\pi i}{q^2 r^2 \left(\frac{1}{n} - i\Theta\right)}\right). \end{aligned}$$

Nun ist

$$\lim_{\tau \rightarrow i\infty} \vartheta(v|\tau) = 1$$

und auf $\xi_{a,q}$ ist $\frac{1}{n} - i\Theta$ klein. Es läßt sich also $\vartheta(v|\tau)$ in erster Annäherung approximieren durch

$$(9) \quad \frac{1}{qr} \sqrt{\frac{\pi}{\frac{1}{n} - i\Theta}} e^{-\frac{\pi^2 \beta^2}{\frac{1}{n} - i\Theta}} S(a, q; b, r),$$

wo

$$S(a, q; b, r) = \sum_{l=0}^{qr-1} e\left(\frac{al^2}{q} + \frac{bl}{r}\right)$$

gesetzt ist.

Falls man die Integrale in (6) in Farey-Bogen bzw. Farey-Segmente aufspaltet und falls man dort überdies $\vartheta(v|\tau)$ durch den Näherungswert (9) ersetzt, so erhält man für $r(n, m)$ den Näherungswert

$$(10) \quad \frac{\pi^{\frac{1}{2}}}{2\pi i} \sum_{a, q, b, r} \left(\frac{S(a, q; b, r)}{qr}\right)^s \int_{\xi_{a,q}} \left(\frac{1}{n} - i\Theta\right)^{-\frac{s}{2}} w^{-n-1} dw \int_{\eta_{b,r}} e^{-\frac{s \pi^2 \beta^2}{\frac{1}{n} - i\Theta}} e^{-2\pi i c m} dv.$$

Dieser Ausdruck wird nun weiter approximiert, indem erstens

$$\int_{\eta_{b,r}} e^{-\frac{s \pi^2 \beta^2}{\frac{1}{n} - i\Theta}} e^{-2\pi i c m} dv = e\left(-\frac{bm}{r}\right) \int_{\eta_{b,r}} e^{-\frac{s \pi^2 \beta^2}{\frac{1}{n} - i\Theta}} e^{-2\pi i \beta m} d\beta,$$

wo über das Intervall

$$-\frac{1}{r(r+r')} \leq \beta \leq \frac{1}{r(r+r')}$$

integriert werden muß, durch

$$e\left(-\frac{bm}{r}\right) \int_{-\infty}^{+\infty} e^{-\frac{s\pi^2\beta^2}{n-i\theta}} e^{-2\pi i\beta m} d\beta = e\left(-\frac{bm}{r}\right) \cdot \frac{1}{\sqrt{\pi s}} \sqrt{\frac{1}{n-i\theta}} \cdot e^{-\frac{m^2}{s}\left(\frac{1}{n-i\theta}\right)}$$

ersetzt wird. Dadurch wird (10) ersetzt durch

$$\frac{\pi^{\frac{1}{2}s-\frac{1}{2}}}{\sqrt{s}} \cdot \frac{1}{2\pi i} \sum_{a,q,b,r} \left(\frac{S(a,q;b,r)}{qr}\right)^s e\left(-\frac{an}{q} - \frac{bm}{r}\right) \int_{\frac{1}{2}a,q} \left(\log \frac{1}{W}\right)^{\frac{1-s}{2}} W^{\frac{m^2}{s}-n-1} dW.$$

In diesem Integral werde statt W die neue Integrationsveränderliche

$$U = W^{\frac{1}{s}}$$

eingeführt, die über einen gewissen Kreisbogen integriert werden muß.

Zweitens wird nun in dem Integranden $\left(\log \frac{1}{U}\right)^{\frac{1-s}{2}}$ durch

$$\frac{1}{\Gamma\left(\frac{s-1}{2}\right)} \sum_{\nu=1}^{\infty} \nu^{\frac{s-1}{2}-1} U^{\nu}$$

approximiert und der Integrationsweg von dem Kreisbogen auf den ganzen Kreis erweitert. Dadurch erhält man für $r(n, m)$ den neuen Näherungsausdruck

$$\frac{\pi^{\frac{s-1}{2}}}{s^{\frac{s}{2}-1} \Gamma\left(\frac{s-1}{2}\right)} \sum_{a,q,b,r} \left(\frac{S(a,q;b,r)}{qr}\right)^s e\left(-\frac{an}{q} - \frac{bm}{r}\right) \times \\ \times \frac{1}{2\pi i} \int_{\mathfrak{K}} \sum_{\nu=1}^{\infty} \nu^{\frac{s-1}{2}-1} U^{n^2-ns+r-1} dU.$$

Der Integrationsweg \mathfrak{K} ist hier ein Kreis, der den Nullpunkt der komplexen U -Ebene als Mittelpunkt hat. Das Integral ist nur dann verschieden von Null, wenn

$$(11) \quad m^2 < ns$$

ist. Wir schreiben noch

$$(12) \quad \Delta = ns - m^2$$

und erhalten schließlich für $r(n, m)$ folgenden Näherungswert:

$$(13) \quad \varrho(n, m) = \frac{\pi^{\frac{s-1}{2}}}{s^{\frac{s}{2}-1} \Gamma\left(\frac{s-1}{2}\right)} \Delta^{\frac{s-3}{2}} \sum_{a,q,b,r} \left(\frac{S(a,q;b,r)}{qr}\right)^s e\left(-\frac{an}{q} - \frac{bm}{r}\right).$$

Zunächst ist hier nur über diejenigen a, q, b, r zu summieren, für die

$$1 \leq a \leq q \leq N, \quad 1 \leq b \leq r \leq M, \quad (a, q) = (b, r) = 1$$

ist. Wir werden uns aber denken, daß q und r in (13) sämtliche ganze Zahlen durchlaufen und a und b alle durch

$$1 \leq a \leq q, \quad 1 \leq b \leq r, \quad (a, q) = (b, r) = 1$$

definierten ganzen Zahlen.

Weiter werde $\varrho(n, m)$ nur im Falle (11) durch (13) definiert, d. h. wenn $\Delta > 0$ ist. Falls $\Delta = 0$ ist, definieren wir:

$$\varrho(n, m) = 1 \text{ oder } = 0, \text{ je nachdem } m \equiv 0 \pmod{s} \text{ oder } \not\equiv 0 \pmod{s}$$

ist. Im Falle $\Delta = sn - m^2 = 0$ ist also

$$\varrho(n, m) = r(n, m).$$

Denn dann folgt aus (4), daß

$$\sum_{k=1}^s (sx_k - m)^2 = s^2 \sum_{k=1}^s x_k^2 - 2ms \sum_{k=1}^s x_k + sm^2 = s(sn - m^2) = 0,$$

also

$$x_1 = x_2 = \dots = x_s = \frac{m}{s}$$

ist. Wir führen noch die Bezeichnung

$$(14) \quad \sum_{a, q, b, r} \left(\frac{S(a, q; b, r)}{qr} \right)^s e \left(-\frac{an}{q} - \frac{bm}{r} \right) = \mathfrak{S}(n, m)$$

ein und nennen diese unendliche Reihe die *singuläre* Reihe (in Nachfolge von Hardy und Littlewood). Statt (13) kann man dann schreiben:

$$(15) \quad \varrho(n, m) = \frac{\pi^{\frac{s-1}{2}}}{s^{\frac{s}{2}-1} \Gamma\left(\frac{s-1}{2}\right)} \Delta^{\frac{s-3}{2}} \mathfrak{S}(n, m) \quad (\Delta > 0).$$

Es muß schließlich noch betont werden, daß den Ausführungen dieses Paragraphen ein lediglich formaler Sinn beigelegt werden muß. Man kann natürlich die Fehler abschätzen, die bei den verschiedenen oben ausgeführten Annäherungen auftreten. Diese Fehlerabschätzungen werden vom Verfasser in einer späteren Arbeit durchgeführt.

§ 2.

Hilfssätze.

Hilfssatz 1. Es seien a, q, b und r ganze Zahlen und es sei

$$q \geq 1, \quad r \geq 1, \quad (b, r) = 1.$$

Falls dann q nicht durch r teilbar ist, so ist

$$S(a, q; b, r) = 0.$$

Beweis. Falls λ ein vollständiges Restsystem mod q und μ ein vollständiges Restsystem mod r durchläuft, so durchläuft $\lambda + \mu q$ ein vollständiges Restsystem mod qr . Deshalb ist

$$S(a, q; b, r) = \sum_{l \bmod qr} e\left(\frac{al^2}{q} + \frac{bl}{r}\right) = \sum_{\lambda \bmod q} \sum_{\mu \bmod r} e\left(\frac{a\lambda^2}{q} + \frac{b\lambda}{r} + \frac{b\mu q}{r}\right).$$

Die Summe über μ ist 0, wenn nicht r in q aufgeht. Falls aber $q \equiv 0 \pmod{r}$ ist, so gilt

$$(16) \quad S(a, q; b, r) = r \sum_{\lambda \bmod q} e\left(\frac{a\lambda^2}{q} + \frac{b\lambda}{r}\right).$$

Hilfssatz 2. Es seien $a_1, q_1, b_1, r_1, a_2, q_2, b_2, r_2$ ganze Zahlen und es sei

$$q_1 \geq 1, \quad r_1 \geq 1, \quad q_2 \geq 1, \quad r_2 \geq 1, \quad r_1 | q_1, \quad r_2 | q_2,$$

$$(a_1, q_1) = (b_1, r_1) = (a_2, q_2) = (b_2, r_2) = (q_1, q_2) = 1.$$

Dann ist

$$S(a_1, q_1; b_1, r_1) S(a_2, q_2; b_2, r_2) = S(a_1 q_2 + a_2 q_1, q_1 q_2; b_1 r_2 + b_2 r_1, r_1 r_2).$$

Beweis. Es durchlaufe λ_1 ein vollständiges Restsystem mod q_1 und λ_2 ein vollständiges Restsystem mod q_2 . Dann durchläuft $\lambda = q_2 \lambda_1 + q_1 \lambda_2$ ein vollständiges Restsystem mod $q_1 q_2$ und es ist:

$$\lambda^2 \equiv q_2^2 \lambda_1^2 + q_1^2 \lambda_2^2 \pmod{q_1 q_2}.$$

Unter Anwendung von (16) folgt also:

$$\begin{aligned} S(a_1 q_2 + a_2 q_1, q_1 q_2; b_1 r_2 + b_2 r_1, r_1 r_2) \\ &= r_1 r_2 \sum_{\lambda \bmod q_1 q_2} e\left(\frac{(a_1 q_2 + a_2 q_1) \lambda^2}{q_1 q_2} + \frac{(b_1 r_2 + b_2 r_1) \lambda}{r_1 r_2}\right) \\ &= r_1 r_2 \sum_{\lambda_1 \bmod q_1} e\left(\frac{a_1 q_2^2 \lambda_1^2}{q_1} + \frac{b_1 q_2 \lambda_1}{r_1}\right) \sum_{\lambda_2 \bmod q_2} e\left(\frac{a_2 q_1^2 \lambda_2^2}{q_2} + \frac{b_2 q_1 \lambda_2}{r_2}\right) \\ &= r_1 r_2 \sum_{\lambda_1 \bmod q_1} e\left(\frac{a_1 \lambda_1^2}{q_1} + \frac{b_1 \lambda_1}{r_1}\right) \sum_{\lambda_2 \bmod q_2} e\left(\frac{a_2 \lambda_2^2}{q_2} + \frac{b_2 \lambda_2}{r_2}\right), \end{aligned}$$

weil $\lambda_1 q_2$ bzw. $\lambda_2 q_1$ mit λ_1 bzw. λ_2 ein vollständiges Restsystem mod q_1 bzw. q_2 durchläuft. Nochmalige Anwendung von (16) ergibt dann die Behauptung des Hilfssatzes.

Hilfssatz 3. Die zahlentheoretische Funktion $h(n, d)$ der ganzzahligen Argumente n und d sei definiert für jede natürliche Zahl n und jeden Teiler d von n . Für

$$(n_1, n_2) = 1, \quad d_1 | n_1, \quad d_2 | n_2$$

habe die Funktion h die Eigenschaft

$$h(n_1 n_2, d_1 d_2) = h(n_1, d_1) h(n_2, d_2).$$

Dann ist die Funktion

$$g(n) = \sum_{d|n} h(n, d)$$

multiplikativ.

Beweis. Falls $(n_1, n_2) = 1$ ist und d_1 durchläuft sämtliche (positive) Teiler von n_1 und d_2 durchläuft sämtliche Teiler von n_2 , dann durchläuft $d_1 d_2$ sämtliche Teiler von $n_1 n_2$. Es ist also

$$\begin{aligned} g(n_1 n_2) &= \sum_{d|n_1 n_2} h(n_1 n_2, d) = \sum_{d_1|n_1} \sum_{d_2|n_2} h(n_1 n_2, d_1 d_2) \\ &= \sum_{d_1|n_1} \sum_{d_2|n_2} h(n_1, d_1) h(n_2, d_2) = g(n_1) g(n_2). \end{aligned}$$

Hilfssatz 4. Für jede natürliche Zahl q sei

$$(17) \quad B_q = \sum_{r|q} A_q(r, 2)$$

$$(18) \quad A_q(r) = \sum'_{a \bmod q} \sum'_{b \bmod r} \left(\frac{S(a, q; b, r)}{q r} \right)^s e\left(-\frac{na}{q} - \frac{mb}{r}\right),$$

wo s , n und m ganze Zahlen sind und $s \geq 1$ ist. Dann ist die Funktion B_q von q multiplikativ.

Beweis. Falls q_1, q_2, r_1, r_2 natürliche Zahlen sind und

$$(q_1, q_2) = 1, \quad r_1 | q_1, \quad r_2 | q_2$$

ist, so ist nach Hilfssatz 2:

$$\begin{aligned} A_{q_1}(r_1) A_{q_2}(r_2) &= \sum'_{a_1 \bmod q_1} \sum'_{a_2 \bmod q_2} \sum'_{b_1 \bmod r_1} \sum'_{b_2 \bmod r_2} \left(\frac{S(a_1, q_1; b_1, r_1)}{q_1 r_1} \right)^s \left(\frac{S(a_2, q_2; b_2, r_2)}{q_2 r_2} \right)^s \times \\ &\quad \times e\left(-\frac{na_1}{q_1} - \frac{na_2}{q_2} - \frac{mb_1}{r_1} - \frac{mb_2}{r_2}\right) \\ &= \sum'_{a_1 \bmod q_1} \sum'_{a_2 \bmod q_2} \sum'_{b_1 \bmod r_1} \sum'_{b_2 \bmod r_2} \left(\frac{S(a_1 q_2 + a_2 q_1, q_1 q_2; b_1 r_2 + b_2 r_1, r_1 r_2)}{q_1 q_2 r_1 r_2} \right)^s \times \\ &\quad \times e\left(-\frac{n(a_1 q_2 + a_2 q_1)}{q_1 q_2} - \frac{m(b_1 r_2 + b_2 r_1)}{r_1 r_2}\right). \end{aligned}$$

Hier durchläuft $a_1 q_2 + a_2 q_1$ ein teilerfremdes Restsystem mod $q_1 q_2$ und $b_1 r_2 + b_2 r_1$ durchläuft ein teilerfremdes Restsystem mod $r_1 r_2$. Deshalb wird

$$\begin{aligned} A_{q_1}(r_1) A_{q_2}(r_2) &= \sum'_{a \bmod q_1 q_2} \sum'_{b \bmod r_1 r_2} \left(\frac{S(a, q_1 q_2; b, r_1 r_2)}{q_1 q_2 r_1 r_2} \right)^s e\left(-\frac{na}{q_1 q_2} - \frac{mb}{r_1 r_2}\right) \\ &= A_{q_1 q_2}(r_1 r_2). \end{aligned}$$

Die Behauptung folgt jetzt aus (17), indem man Hilfssatz 3 anwendet.

Hilfssatz 5. Die Funktion $\psi(x)$ der reellen Veränderlichen x sei periodisch mit der Periode 1. Es sei q eine natürliche Zahl. Dann ist

$$(19) \quad \sum_{b \bmod q} \psi\left(\frac{b}{q}\right) = \sum_{r|q} \sum'_{b \bmod r} \psi\left(\frac{b}{r}\right).$$

²⁾ Verwechslung dieser B_q mit den (später noch auftretenden) Koeffizienten von Bernoulli ist wohl nicht zu befürchten.

Beweis. In der Summe auf der linken Seite von (19) betrachte man diejenigen Glieder, für die b mit q einen festen Teiler d von q als größten gemeinsamen Teiler hat. Dann wird:

$$\sum_{b \bmod q} \psi\left(\frac{b}{q}\right) = \sum_{d|q} \sum_{\substack{b \bmod q \\ (b,q)=d}} \psi\left(\frac{b}{q}\right).$$

In der inneren Summe werde $c = \frac{b}{d}$ als neue Summationsveränderliche eingeführt. Dann wird

$$\sum_{b \bmod q} \psi\left(\frac{b}{q}\right) = \sum_{d|q} \sum'_{c \bmod \frac{q}{d}} \psi\left(\frac{cd}{q}\right).$$

Die Behauptung folgt jetzt, wenn man hier den komplementären Teiler $d_1 = \frac{q}{d}$ von d als neue Summationsveränderliche statt d einführt.

Hilfssatz 6. Es seien a und b ganze Zahlen und es sei q eine natürliche Zahl. Die Summe

$$S_{a,q,b} = \sum_{x \bmod q} e\left(\frac{ax^2 + bx}{q}\right)$$

ist dann $= 0$, falls

$$b \not\equiv 0 \pmod{(a, q)}$$

ist. Ist aber $b \equiv 0 \pmod{(a, q)}$ und ist überdies q ungerade, dann ist

$$S_{a,q,b} = i^{\frac{(q_1-1)^2}{4}} e\left(-\frac{a_1 b_1^2}{q_1}\right) \left(\frac{a_1}{q_1}\right) \sqrt{q} \sqrt{(a, q)},$$

wo

$$a_1 = \frac{a}{(a, q)}, \quad b_1 = \frac{b}{(a, q)}, \quad q_1 = \frac{q}{(a, q)}$$

und $b_2 \bmod q_1$ bestimmt wird durch

$$b_1 \equiv 2 a_1 b_2 \pmod{q_1}.$$

Falls $q = 2^\lambda$ (λ ganz ≥ 2) und a und b ungerade sind, so ist

$$S_{a,q,b} = 0.$$

Falls $q = 2^\lambda$, a ungerade und b gerade ist, so ist

$$S_{a,q,b} = e\left(-\frac{ab_1^2}{q}\right) S_{a,q,0},$$

wo $b_0 \bmod q$ durch

$$\frac{b}{2} \equiv a b_0 \pmod{q}$$

bestimmt wird.

Beweis. Es werde $(a, q) = d$ gesetzt. Man kann zum Beweise der ersten Behauptung $1 < d < q$ voraussetzen, weil die Behauptung sonst trivial ist.

Falls nun x_1 die Zahlen $0, 1, 2, \dots, q_1 - 1$ und x_2 die Zahlen $0, 1, 2, \dots, d - 1$ durchläuft, so durchläuft $x_1 + q_1 x_2$ die Zahlen $0, 1, 2, \dots, q - 1$. Es ist also

$$S_{a,q,b} = \sum_{x_1=0}^{q_1-1} \sum_{x_2=0}^{d-1} e\left(\frac{a_1 x_1^2}{q_1} + \frac{b x_1}{q} + \frac{b x_2}{d}\right).$$

Die Summe über x_2 ist 0, wenn nicht $b \equiv 0 \pmod{d}$ ist.

Falls aber $b \equiv 0 \pmod{d}$ ist, so wird für ungerades q :

$$\begin{aligned} S_{a,q,b} &= d \sum_{x \bmod q_1} e\left(\frac{a_1 x^2 + b_1 x}{q_1}\right) = d e\left(-a_1 \frac{b_1^2}{q_1}\right) \sum_{x \bmod q_1} e\left(a_1 \frac{x^2 + 2b_1 x + b_1^2}{q_1}\right) \\ &= d e\left(-a_1 \frac{b_1^2}{q_1}\right) S_{a_1, q_1}, \end{aligned}$$

wo

$$S_{a_1, q_1} = S_{a_1, q_1, 0} = \sum_{x \bmod q_1} e\left(\frac{a_1 x^2}{q_1}\right)$$

ist. Die zweite Behauptung folgt jetzt, wenn man noch den bekannten Wert der Gaußschen Summe einsetzt.

Falls $q = 2^\lambda$ ($\lambda \geq 2$) und a und b ungerade sind, so annullieren sich die Glieder

$$x = x_1 \quad \text{und} \quad x = x_1 + 2^{\lambda-1}$$

der Summe $S_{a,q,b}$.

Falls aber b gerade ist, so ist:

$$S_{a,q,b} = \sum_{x \bmod q} e\left(a \frac{x^2 + 2b_0 x + b_0^2}{q}\right) \cdot e\left(-\frac{a b_0^2}{q}\right) = e\left(-\frac{a b_0^2}{q}\right) S_{a,q}.$$

Hilfssatz 7. Für ganzes Δ , natürliches d und natürliches ungerades s werde, falls $d \mid \Delta$ ist, das Symbol $\left\{\frac{\Delta; d}{s}\right\}$ durch

$$(20) \quad \left\{\frac{\Delta; d}{s}\right\} = \left(\frac{\Delta/d}{(s, d)}\right) \left(\frac{d}{s/(s, d)}\right)$$

definiert. Falls dann

$$d_1 \mid \Delta, \quad d_2 \mid \Delta, \quad (d_1, d_2) = 1$$

vorausgesetzt wird, so ist

$$\left\{\frac{\Delta; d_1 d_2}{s}\right\} = \left\{\frac{\Delta; d_1}{s}\right\} \cdot \left\{\frac{\Delta; d_2}{s}\right\}.$$

Beweis. Es sei

$$\Delta = d_1 d_2 \Delta', \quad s = (s, d_1) (s, d_2) s' = (s, d_1 d_2) s'.$$

Dann ist

$$(21) \quad \left\{\frac{\Delta; d_1 d_2}{s}\right\} = \left(\frac{\Delta'}{(s, d_1) (s, d_2)}\right) \left(\frac{d_1 d_2}{s'}\right) = \left(\frac{\Delta'}{(s, d_1)}\right) \left(\frac{d_2}{s'}\right) \cdot \left(\frac{\Delta'}{(s, d_2)}\right) \left(\frac{d_1}{s'}\right).$$

Weil

$$\left(\frac{d_2}{(s, d_1)}\right)^2 = 1$$

ist, ist aber

$$(22) \quad \left(\frac{\Delta'}{(s, d_1)} \right) \left(\frac{d_2}{s'} \right) = \left(\frac{\Delta' d_2}{(s, d_1)} \right) \left(\frac{d_2}{s'(s, d_1)} \right) = \left(\frac{\Delta/d_1}{(s, d_1)} \right) \left(\frac{d_2}{s/(s, d_1)} \right).$$

Ebenso ist auch

$$(23) \quad \left(\frac{\Delta'}{(s, d_2)} \right) \left(\frac{d_1}{s'} \right) = \left(\frac{\Delta/d_2}{(s, d_2)} \right) \left(\frac{d_1}{s/(s, d_2)} \right).$$

Man erhält die Behauptung des Hilfssatzes, wenn man (22) und (23) in (21) einsetzt und die Definition (20) beachtet.

Hilfssatz 8. Für ganzes Δ , natürliches, in Δ aufgehendes N und natürliches ungerades s werde die arithmetische Funktion $\psi_s(\Delta; N)$ durch

$$\psi_s(\Delta; N) = \sum_{d|N} d^{\frac{3-s}{2}} \left\{ \frac{\Delta; d}{s} \right\}$$

definiert. Falls dann

$$N_1 | \Delta, \quad N_2 | \Delta, \quad (N_1, N_2) = 1$$

vorausgesetzt wird, so ist

$$\psi_s(\Delta; N_1 N_2) = \psi_s(\Delta; N_1) \psi_s(\Delta; N_2).$$

Beweis. Wegen Hilfssatz 7 ist

$$\begin{aligned} \psi_s(\Delta; N_1 N_2) &= \sum_{d|N_1 N_2} d^{\frac{3-s}{2}} \left\{ \frac{\Delta; d}{s} \right\} = \sum_{d_1|N_1} \sum_{d_2|N_2} (d_1 d_2)^{\frac{3-s}{2}} \left\{ \frac{\Delta; d_1 d_2}{s} \right\} \\ &= \psi_s(\Delta; N_1) \psi_s(\Delta; N_2). \end{aligned}$$

Bezeichnungen. Für reelles x seien die Funktionen $g_1(x), g_2(x) \dots$ durch folgende Formeln definiert:

$$\begin{aligned} g_{2k}(x) &= \frac{(-1)^{k-1}}{2^{2k-1} \pi^{2k}} \sum_{n=1}^{\infty} \frac{\cos 2\pi n x}{n^{2k}} \quad (k = 1, 2, 3, \dots); \\ g_{2k+1}(x) &= \frac{(-1)^{k-1}}{2^{2k} \pi^{2k+1}} \sum_{n=1}^{\infty} \frac{\sin 2\pi n x}{n^{2k+1}} \quad (k = 0, 1, 2, 3, \dots). \end{aligned}$$

Diese (Bernoullischen) Funktionen sind periodisch mit der Periode eins. Bekanntlich ist:

$$g_1(x) = x - \frac{1}{2} \quad (0 < x < 1);$$

$$\begin{aligned} g_{2k}(x) &= \frac{1}{2^k k!} \left[x^{2k} - \frac{2^k}{2} x^{2k-1} + \sum_{n=1}^k (-1)^{n-1} \binom{2k}{2n} B_n x^{2k-2n} \right] \\ &\quad (k = 1, 2, 3, \dots; 0 \leq x \leq 1); \end{aligned}$$

$$\begin{aligned} g_{2k+1}(x) &= \frac{1}{(2k+1)!} \left[x^{2k+1} - \frac{2^{k+1}}{2} x^{2k} + \sum_{n=1}^k (-1)^{n-1} \binom{2k+1}{2n} B_n x^{2k-2n+1} \right] \\ &\quad (k = 1, 2, 3, \dots; 0 \leq x \leq 1), \end{aligned}$$

wo

$$B_1 = \frac{1}{6}, \quad B_2 = \frac{1}{30}, \quad B_3 = \frac{1}{42}, \dots$$

die Koeffizienten von Bernoulli sind.

Hilfssatz 9. Es sei $2k$ eine gerade natürliche Zahl und s eine natürliche Zahl ≥ 3 . Weiter sei $\chi(n)$ ein eigentlicher Charakter mod s (das für nicht zu s teilerfremdes n gleich 0 definiert wird) mit der Eigenschaft

$$(24) \quad \chi(-1) = +1.$$

Dann ist

$$(25) \quad L(2k, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^{2k}} = (-1)^{k-1} \frac{2^{2k-1} \pi^{2k} \tau(\chi)}{s} \sum_{r=1}^{s-1} \bar{\chi}(r) g_{2k}\left(\frac{r}{s}\right),$$

wo zur Abkürzung

$$\tau(\chi) = \sum_{n=1}^{s-1} \chi(n) e\left(\frac{n}{s}\right)$$

gesetzt ist und $\bar{\chi}$ der zu χ konjugiert-komplexe Charakter ist.

Beweis. Weil $-r$ mit r ein vollständiges Restsystem mod s durchläuft, ist wegen (24):

$$\sum_{r \bmod s} \bar{\chi}(r) \sin \frac{2\pi nr}{s} = \sum_{r \bmod s} \bar{\chi}(-r) \sin \frac{-2\pi nr}{s} = - \sum_{r \bmod s} \bar{\chi}(r) \sin \frac{2\pi nr}{s}$$

und deshalb

$$\sum_{r \bmod s} \bar{\chi}(r) \sin \frac{2\pi nr}{s} = 0.$$

Es ist also

$$\begin{aligned} \sum_{r=1}^{s-1} \bar{\chi}(r) g_{2k}\left(\frac{r}{s}\right) &= \frac{(-1)^{k-1}}{2^{2k-1} \pi^{2k}} \sum_{n=1}^{\infty} \frac{1}{n^{2k}} \sum_{r \bmod s} \bar{\chi}(r) \cos \frac{2\pi nr}{s} \\ &= \frac{(-1)^{k-1}}{2^{2k-1} \pi^{2k}} \sum_{n=1}^{\infty} \frac{1}{n^{2k}} \sum_{r \bmod s} \bar{\chi}(r) e\left(\frac{nr}{s}\right). \end{aligned}$$

Da χ ein eigentlicher Charakter mod s ist, ist die Summe über r nur dann verschieden von Null, wenn $(n, s) = 1$ ist. Letzterenfalls ist aber

$$(26) \quad \sum_{r=1}^{s-1} \bar{\chi}(r) e\left(\frac{nr}{s}\right) = \chi(n) \sum_{r \bmod s} \bar{\chi}(nr) e\left(\frac{nr}{s}\right) = \chi(n) \tau(\bar{\chi}),$$

und dieses Resultat ist auch noch für $(n, s) \neq 1$ richtig. Es wird jetzt:

$$(27) \quad \sum_{r=1}^{s-1} \bar{\chi}(r) g_{2k}\left(\frac{r}{s}\right) = \frac{(-1)^{k-1}}{2^{2k-1} \pi^{2k}} L(2k, \chi) \cdot \tau(\bar{\chi}).$$

Es ist aber wegen (24):

$$(28) \quad \tau(\chi) \tau(\bar{\chi}) = s.$$

Die Behauptung (25) folgt, wenn man (28) in (27) einsetzt.

Hilfssatz 10. *Es sei $2k+1$ eine ungerade natürliche Zahl und s eine natürliche Zahl ≥ 3 . Weiter sei $\chi(n)$ ein eigentlicher Charakter mod s (das für nicht zu s teilerfremdes n gleich 0 definiert wird) mit der Eigenschaft*

$$(29) \quad \chi(-1) = -1.$$

Dann ist

$$(30) \quad L(2k+1, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^{2k+1}} = (-1)^{k-1} \frac{2^{2k} \pi^{2k+1} \tau(\chi)}{is} \sum_{r=1}^{s-1} \bar{\chi}(r) g_{2k+1}\left(\frac{r}{s}\right).$$

Beweis. Weil $-r$ mit r ein vollständiges Restsystem mod s durchläuft, ist wegen (29):

$$\sum_{r \bmod s} \bar{\chi}(r) \cos \frac{2\pi nr}{s} = \sum_{r \bmod s} \bar{\chi}(-r) \cos \frac{2\pi nr}{s} = - \sum_{r \bmod s} \bar{\chi}(r) \cos \frac{2\pi nr}{s}$$

und deshalb

$$\sum_{r \bmod s} \bar{\chi}(r) \cos \frac{2\pi nr}{s} = 0.$$

Es folgt also unter Benutzung von (26):

$$(31) \quad \sum_{r=1}^{s-1} \bar{\chi}(r) g_{2k+1}\left(\frac{r}{s}\right) = \frac{(-1)^{k-1}}{2^{2k} \pi^{2k+1}} \sum_{n=1}^{\infty} \frac{1}{n^{2k+1}} \sum_{r \bmod s} \bar{\chi}(r) \sin \frac{2\pi nr}{s} \\ = \frac{(-1)^{k-1}}{2^{2k} \pi^{2k+1}} \sum_{n=1}^{\infty} \frac{1}{n^{2k+1}} \sum_{r \bmod s} \bar{\chi}(r) e\left(\frac{nr}{s}\right) = \frac{(-1)^{k-1} \tau(\bar{\chi})}{2^{2k} \pi^{2k+1}} L(2k+1, \chi).$$

Wegen (29) ist aber

$$(32) \quad \tau(\chi) \tau(\bar{\chi}) = -s.$$

Die Behauptung (30) folgt, wenn man (32) in (31) einsetzt.

§ 3.

Die singuläre Reihe für ungerades quadratfreies $s \geq 3$.

Die singuläre Reihe, die formal durch die Formel (14) definiert wurde, kann nach Hilfssatz 1 auch folgendermaßen gegeben werden:

$$(33) \quad \mathfrak{S}(n, m) = \sum_{q=1}^{\infty} B_q,$$

wo B_q durch (17) und $A_q(r)$ durch (18) definiert wird.

Um die Konvergenz der singulären Reihe (33) zu untersuchen, schreiben wir für jede Primzahl p :

$$(34) \quad \chi_p = 1 + B_p + B_{p^2} + B_{p^3} + \dots$$

Wir zeigen zunächst, daß die Glieder der Reihe (34) von einer gewissen Stelle an 0 sind, daß es sich also um endliche Summen handelt.

Zunächst werde Hilfssatz 5 angewandt auf die Funktion

$$\psi(x) = q^{-s} \left(\sum_{\lambda=0}^{q-1} e\left(\frac{a\lambda^2}{q} + x\lambda\right) \right)^s e\left(-\frac{na}{q} - mx\right).$$

Es folgt dann

$$\sum_{r|q} \sum'_{b \bmod r} \left(\frac{S(a, q; b, r)}{q^r} \right)^s e\left(-\frac{na}{q} - \frac{mb}{r}\right) = \sum_{b \bmod q} \left(\frac{S_{a, q, b}}{q} \right)^s e\left(-\frac{na + mb}{q}\right).$$

Deshalb ist auch

$$(35) \quad B_q = \sum_{a \bmod q} \sum'_{b \bmod q} \left(\frac{S_{a, b, q}}{q} \right)^s e\left(-\frac{na + mb}{q}\right).$$

Falls q ungerade ist, ist hier nach Hilfssatz 6:

$$S_{a, q, b} = i^{\frac{(q-1)^2}{4}} e\left(-\frac{ax^2}{q}\right) \cdot \left(\frac{a}{q}\right) \sqrt{q},$$

wo $x \bmod q$ durch

$$b \equiv 2ax \pmod{q}$$

bestimmt wird. Setzt man diesen Wert von $S_{a, q, b}$ in (35) ein, so kommt:

$$B_q = i^{\frac{(q-1)^2}{4}} q^{-\frac{1}{2}s} \sum_{a \bmod q} \left(\frac{a}{q}\right)^s \sum_{b \bmod q} e\left(-\frac{as^2 + na + mb}{q}\right).$$

Führt man hier in der inneren Summe x statt b als Summationsbuchstaben ein, so erhält man

$$(36) \quad B_q = i^{\frac{(q-1)^2}{4}} q^{-\frac{1}{2}s} \sum_{a \bmod q} \left(\frac{a}{q}\right)^s e\left(-\frac{an}{q}\right) S_{-as, q, -2ma} \quad (q \text{ ungerade}).$$

Von jetzt an werde dauernd s ungerade, quadratfrei und ≥ 3 vorausgesetzt.

Wir unterscheiden verschiedene Fälle:

1. Es sei $(q, 2s) = 1$. Das S in (36) ist dann nach Hilfssatz 6 gleich

$$i^{\frac{(q-1)^2}{4}} e\left(\frac{as y^2}{q}\right) \cdot \left(\frac{-as}{q}\right) \sqrt{q},$$

wo $y \bmod q$ durch

$$m \equiv sy \pmod{q}$$

bestimmt wird. Wenn man diesen Wert von S in (36) einsetzt und noch in der Summe über a den Buchstaben a durch sa ersetzt, so erhält man nach einer kleinen Umformung (Anwendung des quadratischen Reziprozitätsgesetzes):

$$(37) \quad B_q = \left(\frac{q}{s}\right) q^{\frac{1}{2}-\frac{1}{2}s} c_q(\Delta) \quad (s \text{ ungerade}; (q, 2s) = 1),$$

wo Δ durch (12) gegeben wird und

$$c_q(\Delta) = \sum'_{a \bmod q} e\left(\frac{a\Delta}{q}\right)$$

eine Ramanujansche Summe ist.

Falls auch noch q zu Δ teilerfremd ist, erhält man:

$$(38) \quad B_q = \left(\frac{q}{s}\right) q^{\frac{1}{2} - \frac{1}{2}s} \mu(q) \quad (s \text{ ungerade}; (q, 2s\Delta) = 1).$$

2. Es sei p eine ungerade, in s aufgehende Primzahl und $q = p^\lambda$ (λ ganz ≥ 1). Es sei weiter p^γ die höchste in $\Delta = ns - m^2$ steckende Potenz von p , und

$$s = ps_1, \quad \Delta = p^\gamma \Delta_1$$

(weil s quadratfrei vorausgesetzt ist, ist s_1 nicht durch p teilbar). Wenn man jetzt den aus Hilfssatz 6 folgenden Wert S in (36) einsetzt, so ergibt sich erstens:

$$(39) \quad B_q = 0 \quad \text{falls} \quad \gamma = 0$$

ist. Falls nämlich $\gamma = 0$ ist, ist m nicht durch p teilbar und verschwindet die Summe S in (36) also nach Hilfssatz 6.

Falls aber $\gamma \geq 1$ ist, so ergibt Hilfssatz 6:

$$B_q = i^{\frac{1}{4}s(q-1)^2 + \frac{1}{4}(\frac{q}{p}-1)^2} \left(\frac{-s_1}{q}\right) q^{\frac{1}{2} - \frac{1}{2}s} p^{\frac{1}{2}} \sum'_{a \bmod q} \left(\frac{a}{p}\right) e\left(a \frac{\Delta/p}{q}\right).$$

Die Summe rechts ist nur dann verschieden von 0, wenn $\lambda = \gamma$ ist, und dann ist

$$\sum'_{a \bmod q} \left(\frac{a}{p}\right) e\left(\frac{\Delta/p}{q}\right) = p^{\gamma-1} \sum'_{a \bmod p} \left(\frac{a}{p}\right) e\left(\frac{a\Delta_1}{p}\right) = i^{\frac{(p-1)^2}{4}} \left(\frac{\Delta_1}{p}\right) p^{\gamma - \frac{1}{2}}.$$

Nach einer leichten Rechnung findet man dann (unter Anwendung des quadratischen Reziprozitätsgesetzes):

$$(40) \quad B_q = \left(\frac{-\Delta_1}{p}\right) \left(\frac{q}{s_1}\right) q^{\frac{3-s}{2}} = \left\{\frac{-\Delta_1; p^\gamma}{s}\right\} q^{\frac{3-s}{2}} \quad (\lambda = \gamma \geq 1).$$

3. Es sei $q = 2^\lambda$ (λ ganz ≥ 1). Für $\lambda = 1$ berechnet man sofort:

$$(41) \quad B_2 = (-1)^{m+n}.$$

Es werde deshalb $\lambda \geq 2$ vorausgesetzt. In (35) verschwinden dann nach Hilfssatz 6 die Glieder mit ungeradem b . Auf Grund dieses Hilfssatzes findet man:

$$B_q = \sum'_{a \bmod q} \left(\frac{S_{a,q}}{q}\right)^s \sum_{x \bmod \frac{1}{2}q} e\left(-a \frac{sx^2 + 2mx + n}{q}\right).$$

Bestimmt man $y \bmod q$ durch

$$m = sy \pmod{q},$$

so ist

$$\begin{aligned} \sum_{x \bmod \frac{1}{2}q} e\left(-a \frac{sx^2 + 2mx}{q}\right) &= \frac{1}{2} \sum_{x \bmod q} e\left(-as \frac{x^2 + 2xy + y^2}{q}\right) e\left(\frac{asy^2}{q}\right) \\ &= \frac{1}{2} e\left(\frac{asy^2}{q}\right) S_- \end{aligned}$$

und es wird deshalb:

$$(42) \quad B_q = \frac{1}{2} \sum_{a \bmod q}' \left(\frac{S_{a,q}}{q} \right)^s S_{-as,q} e \left(a \frac{sy^s - n}{q} \right).$$

Nun berechnet man aber leicht, daß hier

$$(43) \quad S_{a,q}^s S_{-as,q} = \left(\frac{2}{s} \right)^{s+1} 2^{\frac{s+1}{2}} q^{\frac{s+1}{2}}$$

ist. Falls man (43) in (42) einsetzt und noch a durch as ersetzt, so erhält man schließlich:

$$(44) \quad B_q = \left(\frac{2}{s} \right)^{s+1} \cdot 2^{\frac{s+1}{2}(1-\lambda)} c_q(\Delta).$$

Auf Grund der erhaltenen Resultate wird jetzt bewiesen:

Satz 1. Die singuläre Reihe (33) ist für ungerades, quadratfreies $s \geq 3$ konvergent und für ungerades quadratfreies $s \geq 5$ sogar absolut konvergent.

Beweis. Die unendliche Reihe

$$(45) \quad \sum_{q=1}^{\infty} \left(\frac{q}{s} \right) \mu(q) q^{\frac{1}{2} - \frac{1}{2}s},$$

in der q alle zu $2s\Delta$ teilerfremden ganzen positiven Zahlen durchläuft, ist bekanntlich für $s \geq 3$ konvergent und für $s > 3$ sogar absolut konvergent (vgl. z. B. E. Landau, Handbuch der Lehre von der Verteilung der Primzahlen, zweiter Band, § 173, S. 631–632). Aus Hilfssatz 4 und (38) geht weiter hervor, daß die singuläre Reihe (33) durch formale Multiplikation der Reihe (45) mit

$$\prod_{p|2s\Delta} \chi_p$$

(wo p sämtliche Primteiler von $2s\Delta$ durchläuft) erhalten werden kann. Nun sind aber die χ_p , wie aus den obigen Rechnungen hervorgeht, endliche Summen. Aus der Konvergenz von (45) folgt also auch die von (33).

Satz 2. Für ungerades quadratfreies $s \geq 3$ und gerades $\Delta > 0$ ist

$$\varrho(n, m) = -2 \left(\frac{-2}{s} \right) N^{\frac{s-3}{2}} s^{-\frac{s-3}{2}} \left(\Gamma \left(\frac{s-1}{2} \right) \right)^{-1} \left(\sum_{r=1}^{s-1} \left(\frac{r}{s} \right) g_{\frac{1}{2}(s-1)} \left(\frac{r}{s} \right) \right)^{-1} \psi_s(-\Delta; N),$$

wo $N = \frac{1}{2} \Delta$ ist und ψ_s die in Hilfssatz 8 definierte zahlentheoretische Funktion ist. Für ungerades Δ ist $\varrho(n, m) = 0$.

Beweis. Aus dem Beweis von Satz 1 folgt:

$$(46) \quad \Theta(n, m) = \prod_{p|2s\Delta} \chi_p \cdot \sum_{\substack{q=1 \\ (q, 2s\Delta)=1}}^{\infty} \left(\frac{q}{s} \right) \mu(q) q^{\frac{1}{2} - \frac{1}{2}s}.$$

Zur Berechnung der χ_p mit $p \mid 2s\Delta$ unterscheiden wir vier Fälle:

1. $p \mid s$, $p \nmid \Delta$. Nach (40) ist

$$(47) \quad \chi_p = 1 + \left\{ \frac{-\Delta; p^\gamma}{s} \right\} p^{\gamma \frac{s-s}{2}} = \psi_s(-\Delta; p^\gamma),$$

falls p^γ die genaue, in Δ steckende Potenz von p ist.

2. $p \mid s$, $p \nmid \Delta$, also $\gamma = 0$. Nach (39) ist jetzt

$$(48) \quad \chi_p = 1.$$

3. $p \nmid s$, $p \mid \Delta$, $p \neq 2$. Nach (37) ist:

$$\chi_p = 1 + \sum_{i=1}^{\infty} \left(\frac{p}{s} \right)^i p^{i(\frac{1}{2}-\frac{1}{2})} c_{p^i}(\Delta).$$

Wenn man die bekannten Werte der Ramanujanschen Summen hier einsetzt, so erhält man

$$\chi_p = 1 + \sum_{i=1}^{\gamma} \left(\frac{p}{s} \right)^i p^{\frac{1}{2}i(1-n)} p^{i-1} (p-1) - p^\gamma \left(\frac{p}{s} \right)^{\gamma+1} p^{\frac{1}{2}(\gamma+1)(1-n)}$$

und durch eine leichte Rechnung wird dies:

$$(49) \quad \chi_p = \left(1 - \left(\frac{p}{s} \right) p^{\frac{1}{2}(1-n)} \right) \left(1 + \sum_{i=1}^{\gamma} \left(\frac{p}{s} \right)^i p^{\frac{1}{2}i(3-n)} \right) \\ = \left(1 - \left(\frac{p}{s} \right) p^{\frac{1}{2}(1-n)} \right) \psi_s(-\Delta; p^\gamma).$$

4. $p = 2$. Für ungerades Δ ist nach (41) und (44):

$$\chi_2 = 1 + (-1)^{m+n} = 0,$$

weil für ungerades $\Delta = ns - m^2$ die beiden Zahlen m und n verschiedene Parität haben müssen. Wegen (46) ist damit die zweite Behauptung des Satzes bewiesen.

Für gerades Δ wird nach (44) und (41) (weil jetzt m und n dieselbe Parität haben):

$$\chi_2 = 1 + 1 + \sum_{i=2}^{\infty} \left(\frac{2}{s} \right)^{i+1} 2^{\frac{s-1}{2}(1-i)} c_{2^i}(\Delta).$$

Jetzt sei 2^γ die höchste in $N = \frac{1}{2}\Delta$ aufgehende Potenz von 2. Die bekannten Werte der Ramanujanschen Summen ergeben dann:

$$\chi_2 = 2 + \sum_{i=2}^{\gamma+1} \left(\frac{2}{s} \right)^{i+1} 2^{\frac{s-1}{2}(1-i)} \cdot 2^{i-1} - \left(\frac{2}{s} \right)^{\gamma+1} \cdot 2^{-\frac{s-1}{2}(\gamma+1)} \cdot 2^{\gamma+1}$$

und durch eine leichte Rechnung läßt sich dies umformen zu

$$(50) \quad \chi_2 = 2 \left(1 - \left(\frac{2}{s} \right) 2^{\frac{1-s}{2}} \right) \left(1 + \sum_{i=1}^r \left(\frac{2}{s} \right)^i 2^{i \frac{3-s}{2}} \right) \\ = 2 \left(1 - \left(\frac{2}{s} \right) 2^{\frac{1-s}{2}} \right) \psi_s(-\Delta; 2^i).$$

Wenn man jetzt die erhaltenen Werte (47), (48), (49) und (50) in (46) einsetzt, erhält man somit unter Anwendung des Hilfssatzes 8:

$$\Theta(m, n) = 2 \left(1 - \left(\frac{2}{s} \right) 2^{\frac{1-s}{2}} \right) \prod_{\substack{p|d \\ p \neq s}} \left(1 - \left(\frac{p}{s} \right) p^{\frac{1}{2}(1-s)} \right) \times \\ \times \psi_s(-\Delta; N) \sum_{\substack{q=1 \\ (q, 2s)=1}}^{\infty} \left(\frac{q}{s} \right) \mu(q) q^{\frac{1-s}{2}}$$

oder auch

$$(51) \quad \Theta(m, n) = 2 \psi_s(-\Delta; N) \sum_{q=1}^{\infty} \left(\frac{q}{s} \right) \mu(q) q^{\frac{1-s}{2}}.$$

Es ist aber

$$(52) \quad \sum_{q=1}^{\infty} \left(\frac{q}{s} \right) \mu(q) q^{\frac{1-s}{2}} = \frac{1}{L\left(\frac{s-1}{2}, \chi\right)},$$

wo χ der quadratische Charakter

$$\chi(q) = \left(\frac{q}{s} \right)$$

mod s ist. Es läßt sich nun (52) auf Grund von den Hilfssätzen 9 und 10 auswerten. Falls erstens $s \equiv 1 \pmod{4}$ ist, so ist nach Hilfssatz 9 mit $k = \frac{s-1}{4}$ [weil dann $\chi(-1) = \left(\frac{-1}{s} \right) = +1$ ist]:

$$(53) \quad L\left(\frac{s-1}{2}, \chi\right) = (-1)^{\frac{s-5}{4}} 2^{\frac{s-3}{2}} \pi^{\frac{s-1}{2}} s^{-\frac{1}{2}} \sum_{r=1}^{s-1} \left(\frac{r}{s} \right) g_{\frac{1}{2}(s-1)}\left(\frac{r}{s}\right) \quad (s \equiv 1 \pmod{4}),$$

da die Gaußsche Summe

$$\tau(\chi) = \sum_{n=1}^{s-1} \left(\frac{n}{s} \right) e\left(\frac{n}{s}\right)$$

für $s \equiv 1 \pmod{4}$ den Wert \sqrt{s} hat.

Ist zweitens $s \equiv 3 \pmod{4}$, so ist nach Hilfssatz 10 mit $k = \frac{s-3}{4}$ [weil dann $\chi(-1) = \left(\frac{-1}{s} \right) = -1$ ist]:

$$(54) \quad L\left(\frac{s-1}{2}, \chi\right) = (-1)^{\frac{s-7}{4}} 2^{\frac{s-3}{2}} \pi^{\frac{s-1}{2}} s^{-\frac{1}{2}} \sum_{r=1}^{s-1} \left(\frac{r}{s} \right) g_{\frac{1}{2}(s-1)}\left(\frac{r}{s}\right) \quad (s \equiv 3 \pmod{4}),$$

weil die Gaußsche Summe $\tau(\chi)$ im Falle $s \equiv 3 \pmod{4}$ den Wert $i\sqrt{s}$ hat.

Die beiden Formeln (53) und (54) lassen sich zusammenfassen zu

$$(55) \quad L\left(\frac{s-1}{2}, \chi\right) = -\left(\frac{-2}{s}\right) 2^{\frac{s-3}{2}} \pi^{\frac{s-1}{2}} s^{-\frac{1}{2}} \sum_{r=1}^{s-1} \left(\frac{r}{s}\right) g_{\frac{1}{2}}(s-1)\left(\frac{r}{s}\right).$$

Die Behauptung des Satzes 2 folgt jetzt aus (15), wenn man (55) und (52) in (51) einsetzt.

§ 4.

Die Funktion $\Theta_s(v|\tau)$.

Wir betrachten jetzt die Funktion

$$(56) \quad \Theta(v|\tau) = \Theta_s(v|\tau) = \sum_{n=-\infty}^{+\infty} \sum_{\substack{m=-\infty \\ m^2 \leq sn}}^{+\infty} \varrho(n, m) w^n e^{2\pi i v m},$$

wo wieder $w = e^{\pi i \tau}$ ist und $\varrho(n, m)$ den in Satz 2 angegebenen Wert hat, falls $m^2 < sn$ ist, während $\varrho(n, m)$ im Falle $m^2 = sn$ bereits in § 1 definiert wurde.

Falls

$$\tau = \tau_1 + i\tau_2, \quad v = v_1 + iv_2 \quad (\tau_1, \tau_2, v_1, v_2 \text{ reell})$$

gesetzt wird, so hat die unendliche Reihe (56) (wie leicht aus Satz 2 gefolgert wird) die Reihe

$$\sum_{n=-\infty}^{+\infty} n^{\frac{s-2}{2}} + s e^{-n\pi\tau_2 + 2\pi|v_2|\sqrt{sn}}$$

als Majorante (bis auf eine nur von s und s abhängige Konstante). Es ist deshalb $\Theta(v|\tau)$ eine analytische Funktion von v und τ , die als Funktion von τ in der oberen Halbebene regulär ist und als Funktion von v eine ganze Funktion ist. Ähnliche Betrachtungen gelten für alle in diesem Paragraphen durch Reihen definierten Funktionen.

Um jetzt die Invarianzeigenschaften der Funktion $\Theta(v|\tau)$ zu ergründen, wird weiterhin ein vielfacher Gebrauch gemacht von den Eisensteinschen Reihen

$$(57) \quad G_k(\tau; \lambda, \mu, s) = \sum_{\substack{n_1 \equiv \lambda(s) \\ n_2 \equiv \mu(s)}}'' \frac{1}{(n_1\tau + n_2)^k},$$

in denen n_1, n_2 alle ganzen rationalen Zahlen, die den angegebenen Bedingungen genügen, durchlaufen sollen und der Akzent den Ausschluß des Wertepaares $n_1 = n_2 = 0$ bedeutet. Wie oben ist dabei τ eine komplexe Variable mit positivem Imaginärteil und es sind k, λ, μ, s ganze rationale Zahlen, und zwar soll immer $k \geq 1, s \geq 3$ sein. Die Reihen (57) sind ausführlich von

Hecke untersucht worden³⁾. Von diesen Heckeschen Untersuchungen zitieren wir einige Resultate, die im folgenden gebraucht werden.

Für $k > 2$ sind die Doppelreihen (57) absolut konvergent. Sie sind dann analytische Funktionen von τ , die in der oberen Halbebene $\Im(\tau) > 0$ regulär sind. Für $k = 1$ und $k = 2$ sind die Reihen (57) nicht mehr absolut konvergent. Unter $G_1(\tau; \lambda, \mu, s)$ bzw. $G_2(\tau; \lambda, \mu, s)$ versteht man nach Hecke (l. c. S. 207 und 213) den Wert der analytischen Funktion von z :

$$\sum'_{\substack{n_1 \equiv \lambda(s) \\ n_2 \equiv \mu(s)}} \frac{1}{(n_1 \tau + n_2) |n_1 \tau + n_2|^z}$$

bzw.

$$\sum'_{\substack{n_1 \equiv \lambda(s) \\ n_2 \equiv \mu(s)}} \frac{1}{(n_1 \tau + n_2)^2 |n_1 \tau + n_2|^z}$$

für $z = 0$. Bei Modulsstitutionen in τ verhalten sich dann die sämtlichen G_k mit $k \geq 1$ wie folgt:

$$(58) \quad G_k\left(\frac{a\tau + b}{c\tau + d}; \lambda, \mu, s\right) = (c\tau + d)^k G_k(\tau; a\lambda + c\mu, b\lambda + d\mu, s)$$

(wo also a, b, c, d ganz rational sind und $ad - bc = 1$ ist).

Es gilt für ein G_k mit $k > 2$ folgende Entwicklung (l. c. S. 201):

$$(59) \quad G_k(\tau; \lambda, \mu, s) = \delta\left(\frac{\lambda}{s}\right) \sum'_{m \equiv \mu(s)} \frac{1}{m^k} + \\ + \frac{(-2\pi i)^k}{s^k (k-1)!} \sum_{\substack{m m_1 > 0 \\ m_1 \equiv \lambda(s)}} m^{k-1} \operatorname{sgn} m \cdot e\left(\frac{\mu m}{s} + \frac{m m_1 \tau}{s}\right),$$

wo $\delta(x) = 1$ oder $= 0$, je nachdem x ganz rational oder nicht ganz rational ist.

Falls $k = 2$ ist, kommt auf der rechten Seite von (59) noch ein Glied hinzu, daß nicht-analytisch in τ ist, und zwar

$$(60) \quad -\frac{2\pi i}{s^2 (\tau - \bar{\tau})},$$

wo $\bar{\tau}$ die konjugiert imaginäre Größe zu τ ist (l. c. S. 207).

Für $k = 1$ gilt folgende Entwicklung (l. c. S. 213):

$$(61) \quad G_1(\tau; \lambda, \mu, s) = C(\lambda, \mu, s) - \frac{2\pi i}{s} \sum_{\substack{m m_1 > 0 \\ m_1 \equiv \lambda(s)}} \operatorname{sgn} m \cdot e\left(\frac{\mu m}{s} + \frac{m m_1 \tau}{s}\right),$$

wo

$$(62) \quad C(\lambda, \mu, s) = \delta\left(\frac{\lambda}{s}\right) \sum'_{m \equiv \mu(s)} \frac{\operatorname{sgn} m}{|m|^{1+z}} \Big|_{z=0} - \frac{\pi i}{s} \sum'_{m \equiv \lambda(s)} \frac{\operatorname{sgn} m}{|m|^z} \Big|_{z=0}$$

ist.

³⁾ E. Hecke, Theorie der Eisensteinschen Reihen höherer Stufe und ihre Anwendung auf Funktionentheorie und Arithmetik. Abh. a. d. math. Sem. der hansischen Univ. 5 (1927), S. 199–224.

Wir führen noch folgende Bezeichnungen ein: Falls v eine ganze rationale Zahl ist, sei

$$\theta, (v|\tau) = \sum_{\substack{m=-\infty \\ m \equiv v(s)}}^{+\infty} \frac{m^2}{w^s} e^{2\pi i v m}.$$

Die arithmetische Funktion $f(\lambda, \mu, v)$ der drei ganzzahligen Argumente λ, μ, v werde = 0 definiert, falls (s, λ) nicht in v aufgeht. Falls aber (s, λ) in v aufgeht, so sei

$$(63) \quad f(\lambda, \mu, v) = \left(\frac{\lambda}{s_1}\right) \left(\frac{2\mu s_1}{(s, \lambda)}\right) i^{\frac{(s, \lambda)-1}{2}} \sqrt{(s, \lambda)} e\left(-\frac{\mu y}{s_1}\right),$$

wo zur Abkürzung

$$s_1 = \frac{s}{(s, \lambda)}$$

gesetzt ist und die ganze rationale Zahl $y \bmod s_1$ durch

$$(64) \quad 2\lambda y \equiv -\frac{v^2}{(s, \lambda)} \pmod{s_1}$$

definiert sei. Weiter sei

$$(65) \quad \varphi, (\tau) = \sum_{\substack{\lambda, \mu \bmod s \\ (\lambda, \mu, s) = 1}} f(\lambda, \mu, v) G_{\frac{1}{2}(s-1)}(\tau; \lambda, \mu, s).$$

Schließlich werde noch zur Abkürzung geschrieben:

$$(66) \quad A_s = -\left(\frac{-2}{s}\right) \cdot 2s^{-\frac{s-3}{2}} \Gamma^{-1}\left(\frac{s-1}{2}\right) \left(\sum_{r=1}^{s-1} \left(\frac{r}{s}\right) g_{\frac{1}{2}(s-1)}\left(\frac{r}{s}\right)\right)^{-1},$$

$$(67) \quad D_s = s^{\frac{s-3}{2}} \Gamma\left(\frac{s-1}{2}\right) (-2\pi i)^{\frac{1-s}{2}} A_s.$$

Wir beweisen zunächst:

Hilfssatz 11. Es ist

$$\sum_{\substack{\lambda, \mu \bmod s \\ (\lambda, \mu, s) = 1}} f(\lambda, \mu, v) = 0.$$

Beweis. Schreibt man zur Abkürzung $(s, \lambda) = \delta$, so wird

$$\sum_{\substack{\lambda, \mu \bmod s \\ (\lambda, \mu, s) = 1}} f(\lambda, \mu, v) = \sum_{\substack{\lambda \bmod s \\ \delta|\lambda}} \left(\frac{\lambda}{s_1}\right) \left(\frac{2s_1}{\delta}\right) i^{\frac{(\delta-1)^2}{4}} \sqrt{\delta} \sum_{\substack{\mu \bmod s \\ (\mu, \delta) = 1}} \left(\frac{\mu}{\delta}\right) e\left(-\frac{\mu y}{s_1}\right).$$

Hier verschwindet aber die Summe über μ identisch in λ und v . Denn setzt man

$$\mu = \mu_1 \delta + \mu_2 s_1,$$

so durchläuft μ gerade die zu δ teilerfremden Restklassen mod s , wenn μ_1 ein volles Restsystem mod s_1 und μ_2 ein teilerfremdes Restsystem mod δ durchläuft. Es ist also

$$\sum_{\substack{\mu \bmod s \\ (u, \delta) = 1}} \left(\frac{\mu}{\delta}\right) e\left(-\frac{\mu y}{s_1}\right) = \left(\frac{s_1}{\delta}\right) \sum_{\mu_2 \bmod \delta} \left(\frac{\mu_2}{\delta}\right) \sum_{\mu_1 \bmod s_1} e\left(-\frac{\mu_1 \delta y}{s_1}\right) = 0.$$

Unter Benutzung der obenerwähnten Heckschen Resultate soll jetzt gezeigt werden, daß für ungerades quadratfreies $s \geq 3$ die Funktion $\Theta(v|\tau)$ dargestellt werden kann als eine lineare Verbindung der Funktionen $\vartheta_v(v|\tau)$, und zwar mit Koeffizienten, die ihrerseits lineare Verbindungen der Funktionen G sind. Genauer gilt:

Satz 3. Für ungerades, quadratfreies $s \geq 5$ ist

$$\Theta(v|\tau) = \frac{1}{2} D_s \sum_{r \bmod s} \varphi_r(\tau) \vartheta_r(v|\tau).$$

Beweis. In der unendlichen Reihe (56) betrachten wir zuerst die Glieder mit $m^2 = sn$. Weil s quadratfrei vorausgesetzt ist, muß für diese Glieder m durch s teilbar und deshalb nach der Definition in § 1:

$$\varrho(n, m) = 1$$

sein. Die Glieder mit $m^2 = sn$ liefern also zu der Funktion $\Theta(v|\tau)$ den Beitrag

$$\sum_{\substack{m=-\infty \\ m \equiv 0(s)}}^{+\infty} w^{\frac{m^2}{s}} e^{2\pi i v m} = \vartheta(s v | s \tau) = \vartheta_0(v|\tau).$$

Schreiben wir also

$$(68) \quad \Theta(v|\tau) = \psi(v|\tau) + \vartheta_0(v|\tau),$$

so wird nach Satz 2 mit der Bezeichnung (66):

$$\psi(v|\tau) = A_s \sum_{\substack{m, n=-\infty \\ m^2 < sn \\ m \equiv n(2)}}^{+\infty} N^{\frac{s-3}{2}} \psi_s(-A; N) w^n e^{2\pi i v m}.$$

Es werde hier

$$\psi_s(-A; N) = \sum_{d|N} d^{\frac{3-s}{2}} \left\{ \frac{A; d}{s} \right\}$$

eingesetzt und es werde die Reihenfolge der Summation geändert. Dann werde statt n eine neue Summationsveränderliche k eingeführt, die durch

$$A = 2N = sn - m^2 = 2dk \quad (k = 1, 2, \dots; m^2 + 2dk \equiv 0(s))$$

definiert ist. Dann wird

$$\psi(v|\tau) = A_s \sum_{m=-\infty}^{+\infty} \frac{m^2}{w^s} e^{2\pi i v m} \sum_{d=1}^{\infty} \left(\frac{d}{s/(s,d)} \right) \sum_{\substack{k=1 \\ 2dk+m^2 \equiv 0(s)}}^{\infty} \left(\frac{-2k}{(s,d)} \right) k^{\frac{s-3}{2}} w^{\frac{2dk}{s}}$$

oder auch

$$(69) \quad \psi(v|\tau) = A_s \sum_{r \bmod s} \vartheta_r(v|\tau) \sum_{d=1}^{\infty} \left(\frac{d}{s/(s,d)} \right) \sum_{\substack{k=1 \\ 2dk+r^2 \equiv 0(s)}}^{\infty} \left(\frac{-2k}{(s,d)} \right) k^{\frac{s-3}{2}} w^{\frac{2dk}{s}}.$$

Andererseits werde die durch (65) definierte Funktion $\varphi_r(\tau)$ nach Potenzen von w entwickelt, indem man in (65) die Entwicklung (59) und den Wert (63) von $f(\lambda, \mu, v)$ einsetzt. Man kann dann die Summation nach μ ausführen, wobei die Summe

$$(70) \quad \sum_{\substack{\mu \bmod s \\ (\mu, \delta) = 1}} \left(\frac{\mu}{\delta} \right) e \left(-\frac{\mu y}{s_1} + \frac{\mu m}{s} \right)$$

[zur Abkürzung ist $(s, \lambda) = \delta$ und $s_1 = \frac{s}{\delta}$ gesetzt] ausgewertet werden muß. Letzteres gelingt leicht, wenn man bemerkt, daß μ gerade die zu δ teilerfremden Restklassen mod s durchläuft, wenn man

$$\mu = \mu_1 \delta + \mu_2 s_1$$

setzt, und wenn dann μ_1 ein volles Restsystem mod s_1 und μ_2 ein teilerfremdes Restsystem mod δ durchläuft. Es zeigt sich dann, daß die Summe (70) nur dann nicht verschwindet, wenn $m \equiv \delta y \pmod{s_1}$ ist, und dann gleich

$$\left(\frac{s_1 m}{\delta} \right) i^{\frac{(\delta-1)^2}{4}} s_1 \sqrt{\delta}$$

ist. Man findet:

$$(71) \quad \begin{aligned} \varphi_r(\tau) &= \delta \left(\frac{v}{s} \right) \sum_{\mu \bmod s} \left(\frac{2\mu}{s} \right) i^{\frac{(s-1)^2}{4}} \sqrt{s} \sum_{\substack{m \equiv u(s) \\ m \equiv \frac{1}{2}(s-1)}} \frac{1}{m^{\frac{1}{2}(s-1)}} + \\ &+ \frac{(-2\pi i)^{\frac{s-1}{2}}}{s^{\frac{s-3}{2}} \Gamma\left(\frac{s-1}{2}\right)} \sum_{\lambda \bmod s} \sum_{\substack{m \bmod s \\ m_1 > 0 \\ (s, \lambda) | v \\ m_1 \equiv \lambda(s) \\ m \equiv \delta y(s_1)}} \left(\frac{\lambda}{s_1} \right) \left(\frac{-2m}{\delta} \right) m^{\frac{s-3}{2}} \operatorname{sgn} m \cdot e \left(\frac{m m_1 \tau}{s} \right) \\ &= \delta \left(\frac{v}{s} \right) \cdot \left(\frac{2}{s} \right) i^{\frac{(s-1)^2}{4}} \sqrt{s} \sum_{m=-\infty}^{+\infty} \left(\frac{m}{s} \right) \frac{1}{m^{\frac{1}{2}(s-1)}} + \\ &+ 2 \frac{A_s}{D_s} \sum_{\substack{m=1 \\ 2m m_1 + v^2 \equiv 0(s)}}^{\infty} \sum_{m_1=1}^{\infty} \left(\frac{m_1}{s/(m_1, s)} \right) \left(\frac{-2m}{(m_1, s)} \right) m^{\frac{s-3}{2}} w^{\frac{2m m_1}{s}} \end{aligned}$$

[denn aus den drei Kongruenzen $m_1 \equiv \lambda \pmod{s}$, $m \equiv \delta y \pmod{s_1}$ und (64) folgt die Kongruenz $2mm_1 + v^2 \equiv 0 \pmod{s}$ und aus den drei Kongruenzen $2mm_1 + v^2 \equiv 0 \pmod{s}$, $m_1 \equiv \lambda \pmod{s}$ und (64) folgt auch umgekehrt, daß $\delta|v$ und $m \equiv \delta y \pmod{s_1}$ ist]. Die erste unendliche Reihe

$$\sum_{m=-\infty}^{+\infty} \left(\frac{m}{s}\right) \frac{1}{m^{\frac{1}{2}(s-1)}} = 2L\left(\frac{s-1}{2}, \chi\right)$$

wird durch (55) gegeben. Eine leichte Rechnung ergibt

$$\delta\left(\frac{v}{s}\right) \cdot \left(\frac{2}{s}\right) i^{\frac{(s-1)^2}{4}} \sqrt{s} \sum_{m=-\infty}^{+\infty} \left(\frac{m}{s}\right) \frac{1}{m^{\frac{1}{2}(s-1)}} = \frac{2}{D_s} \delta\left(\frac{v}{s}\right).$$

Weiter ist die in (71) vorkommende Doppelreihe dieselbe wie in (69) (es ist nur m_1 statt d und m statt k geschrieben). Aus (69) und (71) ergibt sich somit

$$\begin{aligned} \psi(v|\tau) &= A_s \sum_{\tau \bmod s} \vartheta_r(v|\tau) \left[\frac{D_s}{2A_s} \varphi_r(\tau) - \delta\left(\frac{v}{s}\right) \frac{1}{A_s} \right] \\ &= -\vartheta_0(v|\tau) + \frac{1}{2} D_s \sum_{\tau \bmod s} \varphi_r(\tau) \vartheta_r(v|\tau) \end{aligned}$$

und wegen (69) ist Satz 3 damit bewiesen. Zunächst war dabei $\frac{s-1}{2} > 2$, also $s > 5$ vorausgesetzt, weil die Entwicklung (59) nur für $k > 2$ gültig war. Für $s = 5$, d. h. $k = \frac{s-1}{2} = 2$ [in (59)] muß noch das Glied (60) hinzugefügt werden. Dies hat aber keinen Einfluß auf das Endergebnis. Denn zu der rechten Seite von (71) muß

$$-\frac{2\pi i}{s^3(\tau - \bar{\tau})} \sum_{\substack{\lambda, \mu \bmod s \\ (\lambda, \mu, s) = 1}} f(\lambda, \mu, v)$$

hinzugefügt werden und dies ist 0 nach Hilfssatz 11. Es bleibt also Satz 3 auch für $s = 5$ gültig.

Anders verhält es sich aber für $s = 3$. Weil G_k für $k = 1$ die Entwicklung (61) statt (59) hat, muß jetzt in (71) das erste der beiden Glieder auf der rechten Seite durch

$$(72) \quad \sigma = \sum_{\substack{\lambda, \mu \bmod 3 \\ (\lambda, \mu, 3) = 1}} f(\lambda, \mu, v) C(\lambda, \mu, 3)$$

ersetzt werden, wo nach (62):

$$(73) \quad C(\lambda, \mu, 3) = \delta\left(\frac{\lambda}{3}\right) \sum'_{m \equiv \mu(3)} \frac{\operatorname{sgn} m}{|m|^{1+s}} \Big|_{s=0} - \frac{\pi i}{3} \sum'_{m \equiv \lambda(3)} \frac{\operatorname{sgn} m}{|m|^s} \Big|_{s=0}$$

ist. Zur Berechnung von (72) bemerke man, daß nach der Definition (63):

$$\text{falls } \lambda \equiv 0 \pmod{3} \text{ und } \nu \equiv 0 \pmod{3}: f(\lambda, \mu, \nu) = -\left(\frac{\mu}{3}\right) i \sqrt{3};$$

$$\text{falls } \lambda \equiv 0 \pmod{3} \text{ und } \nu \not\equiv 0 \pmod{3}: f(\lambda, \mu, \nu) = 0;$$

$$\text{falls } \lambda \not\equiv 0 \pmod{3}: f(\lambda, \mu, \nu) = \left(\frac{\lambda}{3}\right) e\left(-\frac{\mu \nu}{3}\right).$$

In der letzten Formel ist $2\lambda y \equiv -\nu^2 \pmod{3}$, also $y \equiv \lambda \nu^2 \pmod{3}$ und deshalb ist auch:

$$\text{falls } \lambda \not\equiv 0 \pmod{3}: f(\lambda, \mu, \nu) = \left(\frac{\lambda}{3}\right) e\left(-\frac{\lambda \mu \nu^2}{3}\right).$$

Setzt man diese Werte, sowie auch (73) in (72) ein, so erhält man [wenn man noch bemerkt, daß $\sum_{\mu=1}^2 \left(\frac{\mu}{3}\right) = 1 - 1 = 0$ ist]:

$$\begin{aligned} \sigma = & -i\sqrt{3} \delta\left(\frac{\nu}{3}\right) \sum_{\mu=1}^2 \left(\frac{\mu}{3}\right) \sum_{m \equiv \mu(3)} \frac{\operatorname{sgn} m}{|m|^{1+z}} \Big|_{z=0} + \\ & -\frac{\pi i}{3} \sum_{\lambda=1}^2 \sum_{\mu=0}^2 \left(\frac{\lambda}{3}\right) e\left(-\frac{\lambda \mu \nu^2}{3}\right) \sum_{m \equiv \lambda(3)} \frac{\operatorname{sgn} m}{|m|^z} \Big|_{z=0}. \end{aligned}$$

Da

$$\sum_{\mu=0}^2 e\left(-\frac{\lambda \mu \nu^2}{3}\right)$$

den Wert 0 oder 3 hat, je nachdem $\lambda \nu^2 \not\equiv 0 \pmod{3}$ oder $\equiv 0 \pmod{3}$ ist, so kommt:

$$\sigma = -\delta\left(\frac{\nu}{3}\right) (2i\sqrt{3} L(1) + 2\pi i L(0)),$$

wo $L(z)$ die durch

$$L(z) = \sum_{m=1}^{\infty} \left(\frac{m}{3}\right) \frac{1}{m^z}$$

definierte Funktion ist, die bekanntlich der Funktionalgleichung

$$2^{1-s} \pi^{-s} 3^{s-\frac{1}{2}} \sin \frac{\pi z}{2} \Gamma(z) L(z) = L(1-z)$$

genügt. Für $z=1$ liefert diese Funktionalgleichung:

$$L(0) = \frac{\sqrt{3}}{\pi} L(1)$$

und es wird deshalb

$$\sigma = -\delta\left(\frac{\nu}{3}\right) \cdot 4i\sqrt{3} \cdot L(1).$$

Aus (55) folgt aber für $s=3$:

$$L(1) = -\frac{\pi}{\sqrt{3}} \left(g_1\left(\frac{1}{3}\right) - g_1\left(\frac{2}{3}\right)\right) = \frac{\pi}{3\sqrt{3}}$$

und es wird also schließlich

$$\sigma = -\frac{4\pi i}{3} \cdot \delta\left(\frac{\nu}{3}\right).$$

Bemerkt man noch, daß [wie leicht aus (66) und (67) folgt] $D_3 = \frac{3i}{\pi}$ ist, so liefert die oben beim Beweise des Satzes 3 durchgeführte Rechnung für $s = 3$ den

Satz 4. Für $s = 3$ ist

$$\Theta(\nu|\tau) + \vartheta_0(\nu|\tau) = \frac{1}{2} D_3 \sum_{r \bmod 3} \varphi_r(\tau) \vartheta_r(\nu|\tau).$$

§ 5.

Die Reziprozitätseigenschaft von $f(\lambda, \mu, \nu)$.

Satz 5. Es sei s ungerade, quadratfrei und ≥ 3 . Weiter seien λ, μ und ν ganze Zahlen und es sei $(\lambda, \mu, s) = 1$. Dann ist

$$f(\lambda, \mu, \nu) = i^{\frac{s-1}{2}} s^{-\frac{1}{2}} \sum_{\kappa \bmod s} f(-\mu, \lambda, \kappa) e\left(-\frac{\kappa \nu}{s}\right).$$

Beweis. Wir schreiben

$$(s, \lambda) = \delta_1, \quad (s, \mu) = \delta_2, \quad s = \delta_1 \delta_2 \delta_3.$$

[Weil $(\lambda, \mu, s) = 1$ ist, sind δ_1 und δ_2 teilerfremd und es ist deshalb s durch $\delta_1 \delta_2$ teilbar.] Nach der Definition von $f(-\mu, \lambda, \kappa)$ ist jetzt:

$$(74) \quad \left\{ \begin{aligned} & i^{\frac{s-1}{2}} s^{-\frac{1}{2}} \sum_{\kappa \bmod s} f(-\mu, \lambda, \kappa) e\left(-\frac{\kappa \nu}{s}\right) \\ &= i^{\frac{s-1}{2}} s^{-\frac{1}{2}} \left(\frac{-\mu}{\delta_1 \delta_2}\right) \left(\frac{2\lambda \delta_1 \delta_2}{\delta_3}\right) i^{\frac{(\delta_3-1)^2}{4}} \sqrt{\delta_3} \sum_{\substack{\kappa \bmod s \\ \kappa \equiv 0 \pmod{\delta_2}}} e\left(-\frac{\lambda \kappa}{\delta_1 \delta_2} - \frac{\kappa \nu}{\delta_1 \delta_2 \delta_3}\right) \end{aligned} \right.$$

wo $z \bmod \delta_1 \delta_3$ durch

$$(75) \quad 2\mu z \equiv \frac{\kappa^2}{\delta_2} \pmod{\delta_1 \delta_3}$$

bestimmt wird.

Es soll zuerst die auf der rechten Seite von (74) auftretende Summe über κ (die kurz durch Σ_{κ} angegeben werden soll) ausgewertet werden. Es werde μ_1 derart bestimmt, daß

$$2\mu \mu_1 \equiv 1 \pmod{\delta_1 \delta_3}, \quad (\mu_1, \delta_2) = 1$$

ist. Nach (75) ist dann:

$$z \equiv \frac{\kappa^2}{\delta_2} \mu_1 \pmod{\delta_1 \delta_3}$$

und es wird

$$\Sigma_{\kappa} = \sum_{\substack{\kappa \bmod s \\ \kappa \equiv 0 \pmod{\delta_2}}} e\left(-\frac{\lambda \mu_1 \kappa^2 + \kappa \nu}{s}\right).$$

Diese verallgemeinerte Gaußsche Summe läßt sich leicht auswerten. Falls ν nicht durch $\delta_1 = (s, \lambda)$ teilbar ist, ist sie $= 0$, wie aus Hilfssatz 6 gefolgert wird. In diesem Falle ist also Satz 5 nach der Definition von $f(\lambda, \mu, \nu)$ bereits bewiesen. Wir können also weiterhin ν durch (s, λ) teilbar annehmen und $\nu = \delta_1 \nu_1$ schreiben. Bestimmt man die ganzen Zahlen λ_1 und ν_2 derart, daß

$$\lambda = \delta_1 \lambda_1, \quad \nu_1 \equiv 2 \mu_1 \lambda_1 \delta_2 \nu_2 \pmod{\delta_3}$$

ist, so findet man

$$\Sigma_x = \left(-\frac{\delta_3 \mu_1 \lambda_1}{\delta_3} \right)^{\frac{(\delta_3-1)^2}{4}} e \left(\frac{\delta_3^2 \nu_2^2 \mu_1 \lambda_1}{\delta_3 \delta_3} \right) \cdot \delta_1 \sqrt{\delta_3}.$$

Es werde nun $y \pmod{\delta_2 \delta_3}$ durch (64), also durch

$$2 \lambda y \equiv -\frac{\nu^2}{\delta_1} \pmod{\delta_2 \delta_3}$$

bestimmt. Dann ist auch

$$2 \lambda_1 y \equiv -\nu_1^2 \pmod{\delta_2 \delta_3},$$

also

$$2 \lambda_1 y \equiv -4 \mu_1^2 \lambda_1^2 \delta_2^2 \nu_2^2 \pmod{\delta_3},$$

und weil $2 \lambda_1$ zu δ_3 teilerfremd ist:

$$y \equiv -2 \mu_1^2 \lambda_1 \delta_2^2 \nu_2^2 \pmod{\delta_3}.$$

Dies gibt weiter:

$$\frac{\mu}{\delta_2} y \equiv -2 \mu \mu_1 \cdot \mu_1 \lambda_1 \delta_2 \nu_2^2 \equiv -\mu_1 \lambda_1 \delta_2 \nu_2^2 \pmod{\delta_3}$$

und also

$$\mu y \equiv -\delta_2^2 \nu_2^2 \mu_1 \lambda_1 \pmod{\delta_2 \delta_3},$$

so daß

$$e \left(\frac{\delta_3^2 \nu_2^2 \mu_1 \lambda_1}{\delta_3 \delta_3} \right) = e \left(-\frac{\mu y}{\delta_3 \delta_3} \right)$$

und

$$\Sigma_x = \left(-\frac{\delta_3 \mu_1 \lambda_1}{\delta_3} \right)^{\frac{(\delta_3-1)^2}{4}} e \left(-\frac{\mu y}{\delta_3 \delta_3} \right) \delta_1 \sqrt{\delta_3}$$

wird. Setzt man diesen Wert in (74) ein, so folgt:

$$(76) \left\{ \begin{aligned} & \frac{s-1}{2} s^{-\frac{1}{2}} \sum_{x \pmod{s}} f(-\mu, \lambda, x) e \left(-\frac{x \nu}{s} \right) \\ &= \frac{s-1}{2} + \frac{(\delta_2-1)^2}{4} + \frac{(\delta_3-1)^2}{4} \left(-\frac{\mu}{\delta_1} \right) \left(\frac{2 \lambda \delta_1 \delta_2}{\delta_2} \right) \left(\frac{\delta_3 \mu \mu_1 \lambda_1}{\delta_3} \right) e \left(-\frac{\mu y}{\delta_3 \delta_3} \right) \sqrt{\delta_1}. \end{aligned} \right.$$

Hierin ist

$$\left(\frac{\delta_3 \mu \mu_1 \lambda_1}{\delta_3} \right) = \left(\frac{2}{\delta_3} \right) \left(\frac{\lambda}{\delta_1} \right) \left(\frac{\delta_1 \delta_2}{\delta_3} \right), \quad \left(\frac{2 \lambda \delta_1 \delta_2}{\delta_2} \right) = \left(\frac{2}{\delta_3} \right) \left(\frac{\lambda}{\delta_3} \right) \left(\frac{\delta_1 \delta_2}{\delta_3} \right).$$

Unter Anwendung des quadratischen Reziprozitätsgesetzes folgt also:

$$\begin{aligned} \left(\frac{-\mu}{\delta_1}\right) \left(\frac{2\lambda\delta_1\delta_3}{\delta_2}\right) \left(\frac{\delta_3\mu\mu_1\lambda_3}{\delta_2}\right) &= \left(\frac{\lambda}{\delta_2\delta_3}\right) \left(\frac{\mu}{\delta_1}\right) \cdot \left(\frac{2}{\delta_2\delta_3}\right) \left(\frac{-1}{\delta_1}\right) \cdot \left(\frac{\delta_1\delta_3}{\delta_2}\right) \left(\frac{\delta_1\delta_3}{\delta_2}\right) \\ &= \left(\frac{\lambda}{\delta_2\delta_3}\right) \left(\frac{2\mu}{\delta_1}\right) \cdot \left(\frac{2}{s}\right) \left(\frac{-1}{\delta_1}\right) \cdot \left(\frac{\delta_1\delta_3}{\delta_2}\right) \left(\frac{\delta_1\delta_3}{\delta_2}\right) \\ &= \left(\frac{\lambda}{\delta_2\delta_3}\right) \left(\frac{2\mu\delta_3\delta_2}{\delta_1}\right) \cdot \left(\frac{2}{s}\right) \left(\frac{-1}{\delta_1}\right) \cdot \left(\frac{\delta_3}{\delta_2}\right) \left(\frac{\delta_3}{\delta_2}\right) \times \\ &\quad \times (-1)^{\frac{\delta_2\delta_3-1}{2} \cdot \frac{\delta_1-1}{2}}. \end{aligned}$$

Der Ausdruck (76) wird deshalb

$$\begin{aligned} (77) \quad \left(\frac{\lambda}{\delta_2\delta_3}\right) \left(\frac{2\mu\delta_3\delta_2}{\delta_1}\right) \sqrt{s, \lambda} e \left(-\frac{\mu y}{\delta_2\delta_3}\right) &\cdot i^{\frac{s-1}{2} + \frac{(\delta_2-1)^2}{4} + \frac{(\delta_3-1)^2}{4}} \times \\ &\times \left(\frac{2}{s}\right) \left(\frac{-1}{\delta_1}\right) \cdot (-1)^{\frac{\delta_2-1}{2} \cdot \frac{\delta_3-1}{2} + \frac{\delta_2\delta_3-1}{2} \cdot \frac{\delta_1-1}{2}} \end{aligned}$$

Hierin ist noch

$$i^{\frac{s-1}{2}} \left(\frac{2}{s}\right) = i^{-\frac{(s-1)^2}{4}}$$

Für den Ausdruck

$$(78) \quad i^{\frac{s-1}{2} + \frac{(\delta_2-1)^2}{4} + \frac{(\delta_3-1)^2}{4}} \left(\frac{2}{s}\right) \left(\frac{-1}{\delta_1}\right) (-1)^{\frac{\delta_2-1}{2} \cdot \frac{\delta_3-1}{2} + \frac{\delta_2\delta_3-1}{2} \cdot \frac{\delta_1-1}{2}}$$

findet man also die folgenden Werte:

falls $\delta_1 \equiv 1, \delta_2 \equiv 1, \delta_3 \equiv 1 \pmod{4}$ ist, den Wert 1;

falls $\delta_1 \equiv 1, \delta_2 \equiv 1, \delta_3 \equiv 3 \pmod{4}$ ist, den Wert $i^{-1} i = 1$;

falls $\delta_1 \equiv 1, \delta_2 \equiv 3, \delta_3 \equiv 1 \pmod{4}$ ist, den Wert $i^{-1} i = 1$;

falls $\delta_1 \equiv 1, \delta_2 \equiv 3, \delta_3 \equiv 3 \pmod{4}$ ist, den Wert $i^2 \cdot (-1) = 1$;

falls $\delta_1 \equiv 3, \delta_2 \equiv 1, \delta_3 \equiv 1 \pmod{4}$ ist, den Wert $i^{-1} (-1) = i$;

falls $\delta_1 \equiv 3, \delta_2 \equiv 1, \delta_3 \equiv 3 \pmod{4}$ ist, den Wert $i(-1)(-1) = i$;

falls $\delta_1 \equiv 3, \delta_2 \equiv 3, \delta_3 \equiv 1 \pmod{4}$ ist, den Wert $i(-1)(-1) = i$;

falls $\delta_1 \equiv 3, \delta_2 \equiv 3, \delta_3 \equiv 3 \pmod{4}$ ist, den Wert $i^{-1} \cdot i \cdot i(-1)(-1) = i$.

Der Ausdruck (78) ist also gleich 1, falls $\delta_1 \equiv 1 \pmod{4}$, und gleich i , falls $\delta_1 \equiv 3 \pmod{4}$ ist, d. h. gleich

$$i^{\frac{(\delta_1-1)^2}{4}}$$

Für (77) findet man also

$$\left(\frac{\lambda}{\delta_2\delta_3}\right) \left(\frac{2\mu\delta_3\delta_2}{\delta_1}\right) \sqrt{s, \lambda} i^{\frac{(\delta_1-1)^2}{4}} e \left(-\frac{\mu y}{\delta_2\delta_3}\right),$$

d. h. $f(\lambda, \mu, v)$. Damit ist Satz 5 bewiesen.

§ 6.

Die Invarianzeigenschaften von $\Theta(v|\tau)$.

In diesem Paragraphen wird wieder dauernd s ungerade, quadratfrei und ≥ 3 vorausgesetzt.

Es wird gezeigt, daß die Funktion $\Theta(v|\tau)$ ähnliche Invarianzeigenschaften besitzt, wie die gewöhnlichen Theta-Reihen $\vartheta(v|\tau)$, d. h. daß sie in bezug auf v (bei konstantem τ) eine quasi-elliptische Funktion ist, und daß sie sich weiter in bezug auf Modulsstitutionen in τ auch in einer ähnlichen einfachen Weise reproduziert wie $\vartheta(v|\tau)$.

Satz 6. Es ist

- a) $\Theta(v+1|\tau) = \Theta(v|\tau)$;
- b) $\Theta(v+\frac{1}{2}|\tau) = \Theta(v|\tau+1)$;
- c) $\Theta(v+\tau|\tau) = e^{-2s\pi iv} e^{-s\pi i\tau} \Theta(v|\tau)$.

Beweis. a) Folgt unmittelbar aus (56).

b) Aus (56) folgt:

$$\Theta(v+\frac{1}{2}|\tau) = \sum_{\substack{m, n = -\infty \\ m^2 \leq sn}}^{+\infty} \varrho(n, m) w^n e^{2\pi i m} e^{\pi i m}$$

und

$$\Theta(v|\tau+1) = \sum_{\substack{m, n = -\infty \\ m^2 \leq sn}}^{+\infty} \varrho(n, m) w^n e^{\pi i n} e^{2\pi i m}.$$

Falls aber $\varrho(n, m) \neq 0$ ist, ist $m \equiv n \pmod{2}$ (nach Satz 2). Es ist also $e^{\pi i n} = e^{\pi i m}$, womit die Behauptung b) bewiesen ist.

c) Es ist

$$\begin{aligned} \vartheta_*(v+\tau|\tau) &= \sum_{\substack{m = -\infty \\ m \equiv v(s)}}^{+\infty} w^{\frac{m^2}{s} + 2m} e^{2\pi i m} = w^{-s} e^{-2\pi i v s} \sum_{\substack{m = -\infty \\ m \equiv v(s)}}^{+\infty} w^{\frac{(m+s)^2}{s}} e^{2\pi i v(m+s)} \\ &= e^{-2s\pi i v} e^{-s\pi i\tau} \vartheta_*(v|\tau). \end{aligned}$$

Die Behauptung ergibt sich jetzt aus den Sätzen 3 und 4.

Satz 7. Es ist

- a) $\Theta(v|\tau+1) = \Theta(v+\frac{1}{2}|\tau)$;
- b) $\Theta(v|\tau+2) = \Theta(v|\tau)$.

Beweis. Die erste Behauptung ist dieselbe wie die des Satzes 6b. Durch zweimalige Anwendung dieser ersten Behauptung ergibt sich dann nach Satz 6a die zweite Behauptung.

Hilfssatz 12. Es ist

$$\vartheta_*\left(\frac{v}{\tau} \middle| -\frac{1}{\tau}\right) = \left(\frac{-i\tau}{s}\right)^{\frac{1}{2}} e^{\frac{s\pi i v^2}{\tau}} \sum_{x \bmod s} e\left(-\frac{xv}{s}\right) \vartheta_x(v|\tau).$$

Beweis. Für die Funktion (5) gilt bekanntlich

$$\vartheta\left(\frac{v}{\tau}\middle|\middle|-\frac{1}{\tau}\right) = (-i\tau)^{\frac{1}{2}} e^{\frac{1}{2}\frac{\pi i v^2}{\tau}} \vartheta(v|\tau).$$

Weiter zeigt man sofort, daß

$$\vartheta_r(v|\tau) = w^{\frac{r^2}{2}} e^{2\pi i v r} \vartheta(sv + r|\tau/s)$$

ist. Deshalb wird

$$\begin{aligned} \vartheta_r\left(\frac{v}{\tau}\middle|\middle|-\frac{1}{\tau}\right) &= e^{-\frac{r^2}{2}\frac{\pi i}{\tau}} e^{\frac{2\pi i v r}{\tau}} \vartheta\left(\frac{sv+r}{\tau}\middle|\middle|-\frac{s}{\tau}\right) = e^{-\frac{r^2}{2}\frac{\pi i}{\tau}} e^{\frac{2\pi i v r}{\tau}} \vartheta\left(\frac{v-\frac{r}{s}}{\tau/s}\middle|\middle|-\frac{1}{\tau/s}\right) \\ &= \left(-i\frac{\tau}{s}\right)^{\frac{1}{2}} e^{-\frac{r^2}{2}\frac{\pi i}{\tau}} e^{\frac{2\pi i v r}{\tau}} e^{\frac{\pi i}{\tau}} \left(v-\frac{r}{s}\right)^2 \vartheta\left(v-\frac{r}{s}\middle|\middle|\frac{\tau}{s}\right) \\ &= \left(-i\frac{\tau}{s}\right)^{\frac{1}{2}} e^{\frac{\pi i v^2}{\tau}} \sum_{m=-\infty}^{+\infty} w^{\frac{m^2}{2}} e^{2\pi i m\left(r-\frac{r}{s}\right)} \\ &= \left(-i\frac{\tau}{s}\right)^{\frac{1}{2}} e^{\frac{\pi i v^2}{\tau}} \sum_{x \bmod s} e^{-\frac{2\pi i r x}{s}} \sum_{\substack{m=-\infty \\ m \equiv x(s)}}^{+\infty} w^{\frac{m^2}{2}} e^{2\pi i m r} \\ &= \left(-i\frac{\tau}{s}\right)^{\frac{1}{2}} e^{\frac{\pi i v^2}{\tau}} \sum_{x \bmod s} e\left(-\frac{xv}{s}\right) \vartheta_x(v|\tau). \end{aligned}$$

Satz 8. Für $s \geq 5$ ist

$$\Theta\left(\frac{v}{\tau}\middle|\middle|-\frac{1}{\tau}\right) = (-i\tau)^{\frac{1}{2}} e^{\frac{1}{2}\frac{\pi i v^2}{\tau}} \Theta(v|\tau).$$

Beweis. Nach Satz 3 ist

$$\Theta\left(\frac{v}{\tau}\middle|\middle|-\frac{1}{\tau}\right) = \frac{1}{2} D_s \sum_{r \bmod s} \varphi_r\left(-\frac{1}{\tau}\right) \vartheta_r\left(\frac{v}{\tau}\middle|\middle|-\frac{1}{\tau}\right),$$

und nach Hilfssatz 12 ist also

$$(79) \quad \Theta\left(\frac{v}{\tau}\middle|\middle|-\frac{1}{\tau}\right) = \frac{1}{2} D_s \left(-i\frac{\tau}{s}\right)^{\frac{1}{2}} e^{\frac{1}{2}\frac{\pi i v^2}{\tau}} \sum_{x \bmod s} \vartheta_x(v|\tau) \sum_{r \bmod s} e\left(-\frac{xv}{s}\right) \varphi_r\left(-\frac{1}{\tau}\right).$$

Nach (58) ist weiter (der Index $k = \frac{s-1}{2}$ der G_k wird hier weggelassen)

$$G\left(-\frac{1}{\tau}\middle|\lambda, \mu, s\right) = \tau^{\frac{s-1}{2}} G(\tau; \mu, -\lambda, s)$$

und deshalb wird auf Grund der Definition (65)

$$\begin{aligned} \sum_{r \bmod s} e\left(-\frac{xv}{s}\right) \varphi_r\left(-\frac{1}{\tau}\right) &= \tau^{\frac{s-1}{2}} \sum_{\substack{\lambda, \mu \bmod s \\ (\lambda, \mu, s)=1}} \sum_{r \bmod s} f(\lambda, \mu, r) G(\tau; \mu, -\lambda, s) e\left(-\frac{xv}{s}\right) \\ &= \tau^{\frac{s-1}{2}} \sum_{\substack{\lambda, \mu \bmod s \\ (\lambda, \mu, s)=1}} G(\tau; \lambda, \mu, s) \sum_{r \bmod s} f(-\mu, \lambda, r) e\left(-\frac{xv}{s}\right). \end{aligned}$$

Nach Satz 5 wird dies

$$\frac{s-1}{\tau^2} \cdot \frac{s-1}{s^2} \frac{1}{s^2} \sum_{\substack{\lambda, \mu \bmod s \\ (\lambda, \mu, s) = 1}} f(\lambda, \mu, \kappa) G(\tau; \lambda, \mu, s).$$

Setzt man dies in (79) ein, so wird also

$$\begin{aligned} \Theta\left(\frac{v}{\tau} \middle| -\frac{1}{\tau}\right) &= \frac{1}{2} D_s (-i\tau)^{\frac{1}{2}s} e^{\frac{s\pi i v^2}{\tau}} \sum_{x \bmod s} \vartheta_x(v|\tau) \sum_{\substack{\lambda, \mu \bmod s \\ (\lambda, \mu, s) = 1}} f(\lambda, \mu, \kappa) G(\tau; \lambda, \mu, s) \\ &= (-i\tau)^{\frac{1}{2}s} e^{\frac{s\pi i v^2}{\tau}} \Theta(v|\tau). \end{aligned}$$

§ 7.

Der Nullpunkt $v = \frac{1}{2} + \frac{1}{2}\tau$ von $\Theta(v|\tau)$.

In diesem Paragraphen wird dauernd s ungerade, quadratfrei und ≥ 5 vorausgesetzt.

Satz 9. $\Theta(\frac{1}{2} + \frac{1}{2}\tau|\tau) = 0$.

Beweis. Setzt man

$$\lambda_+(\tau) = \sum_{\substack{m = -\frac{s-1}{2} \\ m \equiv \tau(s)}}^s (-1)^m w^{\frac{m(m+s)}{s}}$$

so ist

$$\vartheta_+(\frac{1}{2} + \frac{1}{2}\tau|\tau) = \lambda_+(\tau) - \lambda_{-s}(\tau).$$

Nach Satz 3 ist also

$$(80) \quad \Theta(\frac{1}{2} + \frac{1}{2}\tau|\tau) = \frac{1}{2} D_s \sum_{r \bmod s} \varphi_r(\tau) (\lambda_r(\tau) - \lambda_{-r}(\tau)).$$

Aus (65), (63), (64) folgt aber, daß $\varphi_r(\tau) = \varphi_{-r}(\tau)$ ist, und dadurch verschwindet (80) identisch in τ .

Es soll gezeigt werden, daß $v = \frac{1}{2} + \frac{1}{2}\tau$ in einigen Fällen eine mehrfache Nullstelle von $\Theta(v|\tau)$ ist. Dazu werden zuerst einige Hilfssätze bewiesen.

Hilfssatz 13. Es sei r eine solche ganze Zahl ≥ 1 , daß für $v = 0, 1, 2, \dots, r-1$ für die v -te Ableitung $\Theta^{(v)}(v|\tau)$ von $\Theta(v|\tau)$ nach v gilt:

$$(81) \quad \Theta^{(v)}(\frac{1}{2} + \frac{1}{2}\tau|\tau) = 0 \quad (v = 0, 1, 2, \dots, r-1)$$

(identisch in τ). Dann hat die Funktion

$$F_r(\tau) = e^{\frac{1}{4}s\pi i \tau} \Theta^{(r)}(\frac{1}{2} + \frac{1}{2}\tau|\tau)$$

folgende Eigenschaften:

$$a) \quad E_r(\tau + 1) = e^{\frac{1}{4}s\pi i} F_r(\tau);$$

$$b) \quad F_r\left(-\frac{1}{\tau}\right) = i^{r+s+2} (-i\tau)^{\frac{s+2r}{2}} F_r(\tau).$$

Beweis. a) [Zum Beweis dieser Behauptung wird die Voraussetzung (81) nicht benutzt.] Durch r -malige Differentiation nach v der Formeln a und b in Satz 6 folgt:

$$(82) \quad \Theta^{(r)}(v+1|\tau) = \Theta^{(r)}(v|\tau), \quad \Theta^{(r)}(v+\frac{1}{2}|\tau) = \Theta^{(r)}(v|\tau+1).$$

Deshalb ist

$$\begin{aligned} F_r(\tau+1) &= e^{\frac{1}{2}\pi i} e^{\frac{1}{2}\pi i \tau} \Theta^{(r)}(1+\frac{1}{2}|\tau+1) = e^{\frac{1}{2}\pi i} e^{\frac{1}{2}\pi i \tau} \Theta^{(r)}(\frac{1}{2}|\tau+1) \\ &= e^{\frac{1}{2}\pi i} e^{\frac{1}{2}\pi i \tau} \Theta^{(r)}(\frac{1}{2}+\frac{1}{2}|\tau) = e^{\frac{1}{2}\pi i} F_r(\tau). \end{aligned}$$

b) Durch r -malige Differentiation nach v der Formel in Satz 8 folgt

$$\Theta^{(r)}\left(\frac{v}{\tau} \middle| -\frac{1}{\tau}\right) = r!(-i\tau)^{\frac{r}{2}} \sum_{k=0}^r \binom{r}{k} \Theta^{(r-k)}(v|\tau) \frac{d^k}{dv^k} e^{\frac{\pi i v^2}{\tau}}$$

Falls man hier $v = \frac{1}{2} + \frac{1}{2}\tau$ einsetzt, so folgt wegen der Voraussetzung (81):

$$\Theta^{(r)}\left(\frac{1}{2} + \frac{1}{2\tau} \middle| -\frac{1}{\tau}\right) = r!(-i\tau)^{\frac{r+2r}{2}} \Theta^{(r)}\left(\frac{1}{2} + \frac{1}{2}|\tau\right) e^{\frac{\pi i}{\tau}\left(\frac{1}{4} + \frac{1}{4}\tau^2 + \frac{1}{2}\tau\right)}$$

oder

$$(83) \quad e^{-\frac{\pi i}{4\tau}} \Theta^{(r)}\left(\frac{1}{2} + \frac{1}{2\tau} \middle| -\frac{1}{\tau}\right) = r!(-i\tau)^{\frac{r+2r}{2}} F_r(\tau).$$

Durch ν -malige Differentiation nach v der Formel c in Satz 6 folgt weiter

$$(84) \quad \Theta^{(\nu)}(v+\tau|\tau) = e^{-\pi i \tau} \sum_{k=0}^{\nu} \binom{\nu}{k} \Theta^{(\nu-k)}(v|\tau) \frac{d^k}{dv^k} e^{-2\pi i \tau v}.$$

Aus (82) und (84) folgt wegen der Voraussetzung (81), daß $\Theta^{(\nu)}(v|\tau)$ für $\nu = 0, 1, 2, \dots, r-1$ die Zahlen $\frac{1}{2} + \frac{1}{2}\tau + g_1 + g_2\tau$ (g_1 und g_2 beliebig ganz) als Nullstellen besitzt. Wählt man in (84)

$$\nu = r, \quad v = \frac{1}{2} - \frac{1}{2}\tau,$$

so folgt jetzt, daß

$$\Theta^{(r)}\left(\frac{1}{2} + \frac{1}{2}\tau|\tau\right) = -\Theta^{(r)}\left(\frac{1}{2} - \frac{1}{2}\tau|\tau\right)$$

ist. Wenn man hier τ durch $-\frac{1}{\tau}$ ersetzt, so kommt:

$$\Theta^{(r)}\left(\frac{1}{2} - \frac{1}{2\tau} \middle| -\frac{1}{\tau}\right) = -\Theta^{(r)}\left(\frac{1}{2} + \frac{1}{2\tau} \middle| -\frac{1}{\tau}\right),$$

und dies, eingesetzt in (83), ergibt die Behauptung b des Hilfssatzes 13.

Hilfssatz 14. Für ganzes $r \geq 1$ sei

$$\gamma_r(\tau) = \frac{F_r(\tau)}{\vartheta^{r+2r}(0|\tau)},$$

wo $F_r(\tau)$ die in Hilfssatz 13 definierte Funktion und $\theta(0|\tau)$ die durch (5) definierte Funktion mit $v=0$ ist. Unter der Voraussetzung (81) gilt dann:

$$\begin{aligned} \text{a) } \gamma_r(\tau+2) &= i^r \gamma_r(\tau); \\ \text{b) } \gamma_r\left(-\frac{1}{\tau}\right) &= i^{r+s+2} \gamma_r(\tau). \end{aligned}$$

Beweis. Die Beweise ergeben sich unmittelbar aus Hilfssatz 13 und den bekannten Formeln

$$\begin{aligned} \theta(0|\tau+2) &= \theta(0|\tau), \\ \theta\left(0\left|-\frac{1}{\tau}\right.\right) &= (-i\tau)^{\frac{1}{2}} \theta(0|\tau). \end{aligned}$$

Hilfssatz 15. Unter der Voraussetzung (81) ist dann und nur dann auch

$$\Theta^{(r)}\left(\frac{1}{2} + \frac{1}{2}\tau|\tau\right) = 0$$

(identisch in τ), falls die Entwicklung von

$$\frac{1}{(2\pi i)^r} w^{-\frac{1}{2}r} \Theta^{(r)}\left(\frac{1}{2} + \frac{1}{2}\tau|\tau\right)$$

nach Potenzen von w keine negativen Potenzen von w enthält.

Beweis. Es werde zunächst $\Theta^{(r)}\left(\frac{1}{2} + \frac{1}{2}\tau|\tau\right)$ nach Potenzen von $w = e^{\pi i \tau}$ entwickelt. Wir schreiben

$$\varphi_s(n) = \varphi_s(-2n; n) = \sum_{d|n} \left(\frac{d}{s|(s, d)}\right) \left(\frac{-2n/d}{(s, d)}\right) d^{\frac{3-s}{2}}.$$

Aus dem Beweise des Satzes 3 [Formel (71)] geht dann hervor, daß

$$\varphi_r(\tau) = \frac{2}{D_s} \delta\left(\frac{r}{s}\right) + \frac{2A_s}{D_s} \sum_{\substack{n=1 \\ 2n \equiv -r^2(s)}}^{\infty} \varphi_s(n) n^{\frac{s-3}{2}} w^{\frac{2n}{s}}$$

ist. Auch findet man ohne Mühe, daß

$$\begin{aligned} \frac{1}{(2\pi i)^r} \theta_r^{(r)}\left(\frac{1}{2} + \frac{1}{2}\tau|\tau\right) &= \sum_{\substack{m=-\frac{s-1}{2} \\ m \equiv r(s)}}^{\infty} (-1)^m m^r w^{\frac{m(m+s)}{2}} + \\ &+ \sum_{\substack{m=-\frac{s-1}{2} \\ m \equiv -r(s)}}^{\infty} (-1)^{m+r-1} (m+s)^r w^{\frac{m(m+s)}{2}} \end{aligned}$$

ist. Wenn man diese Entwicklungen in

$$\Theta^{(r)}\left(\frac{1}{2} + \frac{1}{2}\tau|\tau\right) = \frac{1}{2} D_s \sum_{r \bmod s} \varphi_r(\tau) \theta_r^{(r)}\left(\frac{1}{2} + \frac{1}{2}\tau|\tau\right)$$

einsetzt, so folgt

$$(85) \left\{ \begin{aligned} & \frac{1}{(2\pi i)^r} \Theta^{(r)}\left(\frac{1}{2} + \frac{1}{2}\tau \mid \tau\right) = \sum_{\substack{m=-\frac{s-1}{2} \\ m \equiv 0 \pmod{s}}}^s (-1)^m (m^r + (-1)^{r-1}(m+s)^r) w^{\frac{m(m+s)}{s}} + \\ & + A_s \sum_{v=-\frac{s-1}{2}}^{\frac{s-1}{2}} \sum_{\substack{n=1 \\ 2n \equiv -v^2 \pmod{s}}}^s \sum_{\substack{m=-\frac{s-1}{2} \\ m \equiv v \pmod{s}}}^s (-1)^m \psi_s(n) n^{\frac{s-3}{2}} \times \\ & \times (m^r + (-1)^{r-1}(m+s)^r) w^{\frac{m(m+s)+2n}{s}} \end{aligned} \right.$$

Man betrachte nun die in Hilfssatz 14 definierte Funktion

$$\gamma_r(\tau) = \frac{w^{\frac{s}{4}} \Theta^{(r)}\left(\frac{1}{2} + \frac{1}{2}\tau \mid \tau\right)}{\theta^{s+2r}(0 \mid \tau)}.$$

Weil in (85)

$$\frac{m(m+s)}{s} + \frac{s}{4} = \frac{1}{s} \left(m + \frac{s}{2}\right)^2 > 0$$

ist und

$$\theta(0 \mid \tau) = 1 + 2 \sum_{m=1}^{\infty} w^{m^2}$$

ist, geht nun aus (85) hervor, daß die Potenzreihenentwicklung nach $w = e^{\pi i \tau}$ von $\gamma_r(\tau)$ nur positive Potenzen von w enthält. Deshalb ist

$$(86) \quad \lim_{\tau \rightarrow i\infty} \gamma_r(\tau) = 0 \quad (r = 1, 2, \dots).$$

Die Funktion $(\gamma_r(\tau))^4$ ist weiter nach Hilfssatz 14 eine Modulfunktion für die Gruppe Γ_s aller Moduls substitutionen, welche von den beiden Substitutionen

$$\tau' = \tau + 2 \quad \text{und} \quad \tau' = -\frac{1}{\tau}$$

erzeugt werden. Der Fundamentalbereich dieser Gruppe wird durch die Ungleichungen

$$\Im(\tau) > 0, \quad |\tau| \geq 1, \quad -1 < \Re(\tau) \leq +1$$

definiert und hat die drei Eckpunkte ± 1 und $i\infty$. Auch ist die Funktion $(\gamma_r(\tau))^4$ im Innern dieses Fundamentalbereiches regulär [bekanntlich hat $\theta(0 \mid \tau)$ in der oberen Halbebene keine Nullstellen]. Nach einem bekannten Satz über Modulfunktionen und unter Benutzung von (86) ist es also für das identische Verschwinden von $\gamma_r(\tau)$ und damit auch von $\Theta^{(r)}(\frac{1}{2} + \frac{1}{2}\tau \mid \tau)$ notwendig und hinreichend, daß

$$\lim_{\tau \rightarrow i\infty} \gamma_r\left(\pm 1 - \frac{1}{T}\right) \text{ endlich}$$

ist. Wegen Hilfssatz 14a kann man sich dabei weiter auf die Bedingung

$$(87) \quad \lim_{\tau \rightarrow i\infty} \gamma_r\left(1 - \frac{1}{T}\right) \text{ endlich}$$

beschränken. Es bleibt also nur noch übrig, zu zeigen, daß die Bedingung (87) mit der im Hilfssatz 15 genannten Bedingung äquivalent ist. Dies geht aber hervor aus der Formel

$$\gamma_r \left(1 - \frac{1}{T}\right) = e^{\frac{1}{2} \pi i} i^{r+s+2} \cdot 2^{s+2r} (1 + W^2 + \dots)^{-s-2r} W^{-\frac{1}{2}r} \Theta^{(r)} \left(\frac{1}{2} + \frac{1}{2} T \mid T\right),$$

wo $W = e^{\pi i T}$ gesetzt ist. Der Beweis dieser Formel folgt sofort aus Hilfssatz 13 und der bekannten Formel

$$\vartheta \left(0 \mid 1 - \frac{1}{T}\right) = (-i T)^{\frac{1}{2}} W^{\frac{1}{2}} \sum_{m=-\infty}^{+\infty} W^{m^2+m}.$$

Damit ist Hilfssatz 15 vollständig bewiesen.

§ 8.

Die Fälle $s = 5$ und $s = 7$.

Satz 10. Für $s = 5$ und $s = 7$ ist

$$\Theta_s(v|\tau) = \vartheta^s(v|\tau).$$

Beweis. Es ist

$$\Phi(v|\tau) = \frac{\Theta_s(v|\tau)}{\vartheta^s(v|\tau)}$$

nach Satz 6a und Satz 6c und bekannten Eigenschaften von $\vartheta(v|\tau)$ eine doppeltperiodische Funktion von v mit den Perioden 1 und τ . Wir werden zeigen, daß sie für $s = 5$ und $s = 7$ in ihrem Periodenparallelogramm höchstens einen einfachen Pol besitzt, und betrachten dazu die beiden Fälle $s = 5$ und $s = 7$ getrennt.

a) $s = 5$. Wir betrachten die Entwicklung von

$$\frac{1}{(2\pi i)^r} w^{-\frac{1}{2}r} \Theta_5^{(r)} \left(\frac{1}{2} + \frac{1}{2} \tau \mid \tau\right)$$

nach Potenzen von w . In dieser Entwicklung kommt als negative Potenz von w nur $w^{-\frac{1}{2}r}$ in Frage, und zwar, wie aus (85) hervorgeht, mit dem Koeffizienten

$$A_5 \left[3 \psi_5(3) \{(-1)^r 2^r + (-1)^{r-1} 3^r\} - 2 \psi_5(2) \{(-1)^r 1^r + (-1)^{r-1} 4^r\} \right] + (-1)^{r-1} 5^r.$$

Hier ist, wie man leicht nachrechnet,

$$\psi_5(3) = \frac{2}{3}, \quad \psi_5(2) = \frac{1}{2}, \quad A_5 = 5,$$

und dieser Koeffizient wird also

$$5 \left[2 \{(-1)^r 2^r + (-1)^{r-1} 3^r\} - \{(-1)^r + (-1)^{r-1} 4^r\} \right] + (-1)^{r-1} 5^r.$$

Für $r = 1, 2, 3$ wird dies auffolgend

$$5[2(-2+3) - (-1+4)] + 5 = 0,$$

$$5[2(4-9) - (1-16)] - 25 = 0,$$

$$5[2(-8+27) - (-1+64)] + 125 = 0.$$

Es kommen also in der Entwicklung keine negativen Potenzen von w vor, und aus Satz 9 und Hilfssatz 15 folgt nun, daß $\Theta_5(v|\tau)$ in $v = \frac{1}{2} + \frac{1}{2}\tau$ eine mindestens vierfache Nullstelle hat. Weil $\theta^5(v|\tau)$ in $v = \frac{1}{2} + \frac{1}{2}\tau$ eine genau fünffache Nullstelle hat und weiter in dem Periodenparallelogramm keine Nullstellen aufweist, hat also die elliptische Funktion $\Phi(v|\tau)$ in ihrem Periodenparallelogramm höchstens einen einfachen Pol und ist dort sonst regulär. Nach einem bekannten Satz ist diese elliptische Funktion dann eine Konstante, d. h. es hängt

$$\frac{\Theta_5(v|\tau)}{\theta^5(v|\tau)}$$

nicht von v ab.

b) Eine ähnliche Rechnung gibt dasselbe Resultat auch für $s = 7$. Negative Potenzen von w in der Entwicklung von

$$\frac{1}{(2\pi i)^r} w^{-\frac{1}{2}r} \Theta_7^{(r)}\left(\frac{1}{2} + \frac{1}{2}\tau|\tau\right)$$

können sich in (85) nur ergeben, wenn:

1. $m = -3$, $n = 6$ oder $m = -2$, $n = 5$ oder $m = -1$, $n = 3$ ist in der zweiten Summe in (85) und $m = 0$ ist in der ersten Summe. Man erhält dann $w^{-\frac{1}{2}r}$ mit dem Koeffizienten

$$(-1)^{r-1} 7^r + A_7 [-6^2 \psi_7(6) \{(-1)^r 3^r + (-1)^{r-1} 4^r\} + \\ + 5^2 \psi_7(5) \{(-1)^r 2^r + (-1)^{r-1} 5^r\} - 3^2 \psi_7(3) \{(-1)^r 1^r + (-1)^{r-1} 6^r\}].$$

Hier ist, wie man leicht nachrechnet:

$$A_7 = \frac{7}{8}, \quad 6^2 \psi_7(6) = 40, \quad 5^2 \psi_7(5) = 24, \quad 3^2 \psi_7(3) = 8$$

und der Koeffizient wird also

$$(-1)^{r-1} 7^r + 7 [-5 \{(-1)^r 3^r + (-1)^{r-1} 4^r\} + \\ + 3 \{(-1)^r 2^r + (-1)^{r-1} 5^r\} - \{(-1)^r + (-1)^{r-1} 6^r\}].$$

Für $r = 1, 2, 3, 4, 5$ ist dies auffolgend

$$\begin{aligned} 7 + 7[-5(-3+4) + 3(-2+5) - (-1+6)] &= 0, \\ -7^2 + 7[-5(9-16) + 3(4-25) - (1-36)] &= 0, \\ 7^3 + 7[-5(-27+64) + 3(-8+125) - (-1+216)] &= 0, \\ -7^4 + 7[-5(81-256) + 3(16-625) - (1-1296)] &= 0, \\ 7^5 + 7[-5(-243+1024) + 3(-32+3125) - (-1+7776)] &= 0. \end{aligned}$$

2. Wenn $r = 5$ ist, kommt noch $w^{2-\frac{1}{2}r} = w^{-\frac{1}{2}}$ vor in den Gliedern $m = -3, n = 13; m = -2, n = 12; m = -1, n = 10; m = 0, n = 7; m = 1, n = 3$ von (85), und zwar mit dem Koeffizienten $A_7 = \frac{7}{8}$ multipliziert mit

$$-13^2 \psi_7(13) (-3^5 + 4^5) + 12^2 \psi_7(12) (-2^5 + 5^5) + \\ -10^2 \psi_7(10) (-1^5 + 6^5) + 7^2 \psi_7(7) \cdot 7^5 - 3^2 \psi_7(3) (1 + 8^5).$$

Hierin ist

$$13^2 \psi_7(13) = 168, 12^2 \psi_7(12) = 168, 10^2 \psi_7(10) = 120, 7^2 \psi_7(7) = 48, 3^2 \psi_7(3) = 8.$$

Der Koeffizient wird also

$$7 [-21(4^5 - 3^5) + 21(5^5 - 2^5) - 15(6^5 - 1) + 6 \cdot 7^5 - (1 + 8^5)] = 0.$$

Aus Satz 9 und Hilfssatz 15 folgt nun, daß $\Theta_7(v|\tau)$ in $v = \frac{1}{2} + \frac{1}{2}\tau$ eine mindestens sechsfache Nullstelle hat. In derselben Weise, wie im Falle $s = 5$, folgert man jetzt, daß

$$\frac{\Theta_7(v|\tau)}{\phi^7(v|\tau)}$$

nicht von v abhängt.

Zum Beweise von Satz 10 muß also nur noch gezeigt werden, daß in den Fällen $s = 5$ und $s = 7$:

$$\frac{\Theta_s(0|\tau)}{\phi^s(0|\tau)} = 1$$

ist. Wir betrachten die beiden Fälle wieder getrennt.

a) $s = 5$. Es werde

$$\Phi_5(\tau) = \frac{\Theta_5(0|\tau)}{\phi^5(0|\tau)}$$

gesetzt. Aus Satz 7 und Satz 8 folgt zunächst, daß $\Phi_5(\tau)$ eine zur Gruppe Γ_5 gehörige Modulfunktion ist. Zum Nachweise, daß $\Phi_5(\tau)$ identisch $= 1$ ist, genügt es dann nach einem (oben schon erwähnten) Satz über Modulfunktionen, daß gezeigt wird, daß

$$(88) \quad \lim_{\tau \rightarrow i\infty} \Phi_5(\tau) = 1$$

und

$$(89) \quad \lim_{\tau \rightarrow i\infty} \Phi_5\left(1 - \frac{1}{\tau}\right) = 1$$

ist. Nun ist

$$\Theta_s(0|\tau) = \sum_{\substack{m=-\infty \\ m \equiv 0 \pmod{s}}}^{+\infty} w^{\frac{m^2}{s}} + A_s \sum_{\substack{m \text{ mod } s \\ 2n \equiv -\tau^2 \pmod{s}}} \sum_{\substack{n=1 \\ 2n \equiv -\tau^2 \pmod{s}}}^{\infty} \sum_{\substack{m=-\infty \\ m \equiv \mu \pmod{s}}}^{+\infty} \frac{s-3}{n^2} \psi_s(n) w^{\frac{m^2+2n}{s}}$$

Die zweite Summe enthält nur positive Potenzen von w . Deshalb ist

$$\lim_{\tau \rightarrow i\infty} \Theta_s(0|\tau) = 1.$$

Weil auch

$$\lim_{\tau \rightarrow i\infty} \theta(0|\tau) = 1$$

ist, ist damit (88) bewiesen.

Aus den Sätzen 7 und 8 folgt weiter

$$(90) \quad \Phi_5\left(1 - \frac{1}{T}\right) = \frac{\Theta_5\left(\frac{1}{2}T|T\right)}{\theta^5\left(\frac{1}{2}T|T\right)}.$$

Es ist aber

$$\Theta_s\left(\frac{1}{2}T|T\right) = 2 \sum_{m=0}^{\infty} W^{sm(m+1)} + 2A_s \sum_{\substack{n \equiv 1 \\ 2n \equiv -1^2 (s)}}^{\infty} \sum_{\substack{m = -\frac{s-1}{2} \\ m \equiv n (s)}}^{\infty} n^{\frac{s-3}{2}} \psi_s(n) W^{\frac{sm(m+s)+2n}{s}}$$

In diesen Summen kommen für $s = 5$ keine negativen Potenzen von W vor. Das von W unabhängige Glied erhält man aus dem Glied $m = 0$ der ersten Summe und den Gliedern $m = -2, n = 3$ und $m = -1, n = 2$ der zweiten Summe. Es wird also

$$\lim_{T \rightarrow i\infty} \Theta_5\left(\frac{1}{2}T|T\right) = 2 + 2A_5(3\psi_5(3) + 2\psi_5(2)).$$

Auch ist

$$\theta\left(\frac{1}{2}T|T\right) = 2 \sum_{m=0}^{\infty} W^{m^2+m}$$

und deshalb

$$\lim_{T \rightarrow i\infty} \theta^5\left(\frac{1}{2}T|T\right) = 2^5 = 32.$$

Aus diesen Formeln und (90) folgt nun (89). Damit ist also gezeigt, daß $\Phi_5(\tau)$ identisch gleich 1 ist.

b) $s = 7$. Wie bei $s = 5$ genügt der Nachweis von

$$(91) \quad \lim_{\tau \rightarrow i\infty} \Phi_7(\tau) = 1, \quad \lim_{T \rightarrow i\infty} \Phi_7\left(1 - \frac{1}{T}\right) = 1,$$

wo

$$\Phi_7(\tau) = \frac{\Theta_7(0|\tau)}{\theta^7(0|\tau)}$$

ist. Die erste der Formeln (91) wurde schon oben bei $s = 5$ mitbewiesen. Weiter findet man, wie bei $s = 7$:

$$\begin{aligned} \lim_{T \rightarrow i\infty} \Theta_7\left(\frac{1}{2}T|T\right) &= 2 + 2A_7[6^2\psi_7(6) + 5^2\psi_7(5) + 3^2\psi_7(3)] \\ &= 2 + 2 \cdot \frac{7}{8}[40 + 24 + 8] = 128. \end{aligned}$$

Auch ist

$$\lim_{T \rightarrow i\infty} \theta^7\left(\frac{1}{2}T|T\right) = 2^7 = 128$$

und deshalb

$$\lim_{T \rightarrow i\infty} \Phi_7\left(1 - \frac{1}{T}\right) = \lim_{T \rightarrow i\infty} \frac{\Theta_7\left(\frac{1}{2}T|T\right)}{\theta^7\left(\frac{1}{2}T|T\right)} = \frac{128}{128} = 1,$$

womit auch die zweite der Formeln (91) bewiesen ist. Damit ist der Beweis des Satzes 10 vollständig geführt.

Aus Satz 10 folgen nun nach bekannten Identitätssätzen über Potenzreihen und endliche trigonometrische Summen folgende Sätze:

Hauptsatz I. *Es sei $r_5(n, m)$ die Lösungszahl des diophantischen Gleichungssystems*

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 = n, \quad x_1 + x_2 + x_3 + x_4 + x_5 = m,$$

wo m und n gegebene ganze Zahlen sind. Dann ist $r_5(n, m)$ nur dann $\neq 0$, wenn $m^2 \leq 5n$ und $m \equiv n \pmod{2}$ ist. Falls $m^2 = 5n$ ist, so ist $r_5(n, m) = 1$. Falls aber $m^2 < 5n$ und $m \equiv n \pmod{2}$ ist, so ist

$$r_5(n, m) = 5N \sum_{d|N} \left(\frac{d}{5}\right) \frac{1}{d},$$

falls $N \not\equiv 0 \pmod{5}$ ist, und

$$r_5(n, m) = 5N \left(1 - \left(\frac{N_0}{5}\right) \frac{1}{5^{\beta}}\right) \sum_{d|N_0} \left(\frac{d}{5}\right) \frac{1}{d},$$

falls $N \equiv 0 \pmod{5}$ ist und 5^{β} die höchste in $N = \frac{1}{2}(5n - m^2)$ aufgehende Potenz von 5 und $N = 5^{\beta}N_0$ ist, während die Summen über sämtliche positive Teiler von N bzw. N_0 zu erstrecken sind.

Beweis. Die ersten beiden Behauptungen sind trivial bzw. in § 1 schon bewiesen. Es sei also weiter $m^2 < 5n$ und $m \equiv n \pmod{2}$. Nach Satz 10 ist dann

$$r_5(n, m) = \varrho_5(n, m).$$

Der Wert von $\varrho_5(n, m)$ wird durch Satz 2 gegeben. Es ist, wie man leicht nachrechnet:

$$\sum_{r=1}^4 \left(\frac{r}{5}\right) g_2\left(\frac{r}{5}\right) = \frac{2}{25}.$$

Nach Satz 2 wird deshalb

$$\varrho_5(n, m) = 5N \psi_5(-\Delta; N) = 5N \psi_5(-2N; N).$$

Falls $N \not\equiv 0 \pmod{5}$ ist, so liefert die Definition von $\psi_5(-2N; N)$ in Hilfsatz 8 sofort

$$r_5(n, m) = 5N \psi_5(-2N, N) = 5N \sum_{d|N} \left(\frac{d}{5}\right) \frac{1}{d}.$$

Falls aber $N \equiv 0 \pmod{5}$ ist, so liefert diese Definition:

$$\psi_5(-2N, N) = \sum_{\substack{d|N \\ d \not\equiv 0 \pmod{5}}} \frac{1}{d} \left(\frac{d}{5}\right) - \sum_{d \equiv 0 \pmod{5}} \frac{1}{d} \left(\frac{N/d}{5}\right).$$

In der ersten Summe sind die nicht durch 5 teilbaren Teiler von N gerade die Teiler von N_0 . In der zweiten Summe sind die durch 5 teilbaren Teiler von N ,

für die das Symbol $\left(\frac{N/d}{5}\right)$ nicht verschwindet, gerade die Zahlen $5^\beta d_0$, wo d_0 die Teiler von N_0 durchläuft. Es wird also

$$\psi_5(-2N; N) = \sum_{d|N_0} \frac{1}{d} \left(\frac{d}{5}\right) - \frac{1}{5^\beta} \sum_{d|N_0} \frac{1}{d} \left(\frac{N_0/d}{5}\right).$$

Weil

$$\left(\frac{N_0/d}{5}\right) = \left(\frac{N_0}{5}\right) \left(\frac{d}{5}\right)$$

ist, wird schließlich

$$\psi_5(-2N; N) = \sum_{d|N_0} \frac{1}{d} \left(\frac{d}{5}\right) - \frac{1}{5^\beta} \left(\frac{N_0}{5}\right) \sum_{d|N_0} \frac{1}{d} \left(\frac{d}{5}\right) = \left(1 - \frac{1}{5^\beta} \left(\frac{N_0}{5}\right)\right) \sum_{d|N_0} \frac{1}{d} \left(\frac{d}{5}\right),$$

womit auch die letzte Behauptung des Hauptsatzes I bewiesen ist.

Durch ähnliche Betrachtungen erhält man auch den

Hauptsatz II. *Es sei $r_7(n, m)$ die Lösungszahl des diophantischen Gleichungssystems*

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2 + x_7^2 = n, \quad x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 = m,$$

wo m und n gegebene ganze Zahlen sind.¹ Dann ist $r_7(n, m)$ nur dann $\neq 0$, falls $m^2 \leq 7n$ und $m \equiv n \pmod{2}$ ist. Falls $m^2 = 7n$ ist, ist $r_7(n, m) = 1$. Falls aber $m^2 < 7n$ und $m \equiv n \pmod{2}$ ist, so ist

$$r_7(n, m) = \frac{7}{8} N^2 \sum_{d|N} \left(\frac{d}{7}\right) \frac{1}{d^2},$$

falls $N \not\equiv 0 \pmod{7}$ ist und

$$r_7(n, m) = \frac{7}{8} N^2 \left(1 - \left(\frac{N_0}{7}\right) \frac{1}{7^{2\beta}}\right) \sum_{d|N_0} \left(\frac{d}{7}\right) \frac{1}{d^2},$$

falls $N \equiv 0 \pmod{7}$ ist und 7^β die höchste in $N = \frac{1}{2}(7n - m^2)$ aufgehende Potenz von 7 und $N = 7^\beta N_0$ ist, während die Summen über sämtliche positiven Teiler von N bzw. N_0 zu erstrecken sind.

§ 9.

Der Fall $s = 3$.

Wie schon aus den Sätzen 3 und 4 ersichtlich ist, nimmt der Fall $s = 3$ eine Sonderstellung ein. Es werde die Funktion $\bar{\Theta}_3(v|\tau)$ durch

$$\bar{\Theta}_3(v|\tau) = \frac{1}{2} D_3 \sum_{r \bmod 3} \varphi_r(\tau) \vartheta_r(v|\tau)$$

definiert. Es ist dann $\bar{\Theta}_3(v|\tau)$ nicht mit $\Theta_3(v|\tau)$ identisch, sondern nach Satz 4 ist

$$\bar{\Theta}_3(v|\tau) = \Theta_3(v|\tau) + \vartheta_0(v|\tau).$$

Auch ist Satz 10 für $s = 3$ nicht richtig. Der zu Satz 10 analoge Satz lautet nämlich für $s = 3$:

Satz 11. Es ist

$$\vartheta^3(v|\tau) = \frac{1}{2} \bar{\Theta}_3(v|\tau).$$

Beweis. Die Formel (71) muß für $s = 3$ (wie aus den Betrachtungen des § 4 hervorgeht) durch folgende Formel ersetzt werden:

$$\varphi_r(\tau) = \frac{4}{D_3} \delta\left(\frac{v}{3}\right) + \frac{2A_3}{D_3} \sum_{\substack{n=1 \\ 2n \equiv -v^2(3)}}^{\infty} \psi_3(n) w^{\frac{2n}{3}},$$

wo $A_3 = 6$ und $D_3 = \frac{3i}{\pi}$ ist. Dadurch wird

$$(92) \quad \bar{\Theta}_3(v|\tau) = \frac{3i}{\pi} \sum_{r \bmod 3} \varphi_r(\tau) \vartheta_r(v|\tau) = 2 \sum_{\substack{m=-\infty \\ m \equiv 0(3)}}^{+\infty} w^{\frac{m^2}{3}} e^{2\pi i v m} + \\ + 6 \sum_{\substack{n=1 \\ 2n+m^2 \equiv 0(3)}}^{\infty} \sum_{\substack{m=-\infty \\ m \equiv -v(3)}}^{+\infty} \psi_3(n) w^{\frac{2n+m^2}{3}} e^{2\pi i v m}$$

Hieraus folgt:

$$\bar{\Theta}_3\left(\frac{1}{2} + \frac{1}{2}\tau|\tau\right) = 2 \sum_{\substack{m=-\infty \\ m \equiv 0(3)}}^{+\infty} (-1)^m w^{\frac{m(m+3)}{3}} + 6 \sum_{\substack{n=1 \\ 2n+m^2 \equiv 0(3)}}^{\infty} \sum_{\substack{m=-\infty \\ m \equiv -v(3)}}^{+\infty} (-1)^m \psi_3(n) w^{\frac{m(m+3)+2n}{3}}$$

Dies verschwindet identisch in τ , weil die Glieder $m = m_1$ und $m = -3 - m_1$ sich gegenseitig (in der ersten, wie in der zweiten Summe) annullieren. Also ist

$$\bar{\Theta}_3\left(\frac{1}{2} + \frac{1}{2}\tau|\tau\right) = 0,$$

d. h. die Funktion $\bar{\Theta}_3(v|\tau)$ hat in $v = \frac{1}{2} + \frac{1}{2}\tau$ eine Nullstelle. Daß diese Nullstelle sogar eine zweifache Nullstelle ist, folgt durch den Nachweis von

$$\bar{\Theta}_3'\left(\frac{1}{2} + \frac{1}{2}\tau|\tau\right) = 0.$$

Dieser Nachweis erfolgt durch Betrachtung der Funktionen

$$F_3(\tau) = e^{\frac{1}{2}\pi i \tau} \bar{\Theta}_3'\left(\frac{1}{2} + \frac{1}{2}\tau|\tau\right), \quad \gamma_3(\tau) = \frac{F_3(\tau)}{\theta^5(0|\tau)}.$$

Weil die in Satz 8 für die Funktion Θ bewiesene Funktionalgleichung im Falle $s = 3$ für die Funktion $\bar{\Theta}$ gilt, so ist $(\gamma_3(\tau))^4$ eine zur Gruppe Γ_3 gehörige Modulfunktion, und um das identische Verschwinden von $\gamma_3(\tau)$ [und damit

von $\bar{\Theta}_3\left(\frac{1}{2} + \frac{1}{2} \tau | \tau\right)$ zu zeigen, genügt es also wieder, die Gültigkeit der beiden Formeln

$$\lim_{\tau \rightarrow i\infty} \gamma_3(\tau) = 0 \quad \text{und} \quad \lim_{T \rightarrow i\infty} \gamma_3\left(1 - \frac{1}{T}\right) = 0$$

zu beweisen. Die erste dieser Formeln ergibt sich sofort aus der Gleichung

$$\begin{aligned} \frac{1}{2\pi i} F(\tau) &= 2 \sum_{\substack{m=-\infty \\ m \equiv 0(3)}}^{+\infty} (-1)^m m w^{\frac{1}{3}\left(m + \frac{3}{2}\right)^2} + \\ &+ 6 \sum_{\substack{n=1 \\ 2n+m^2 \equiv 0(3)}}^{\infty} \sum_{m=-\infty}^{+\infty} (-1)^m m \psi_3(n) w^{\frac{2n}{3} + \frac{1}{3}\left(m + \frac{3}{2}\right)^2}, \end{aligned}$$

und die zweite folgt aus den Formeln ($W = e^{\pi i T}$)

$$\begin{aligned} \gamma_3\left(1 - \frac{1}{T}\right) &= - \frac{e^{\frac{3}{2}\pi i} (-i T)^{\frac{5}{2}} F(T)}{2^5 (-i T)^{\frac{5}{2}} W^{\frac{5}{2}} \left(\sum_{m=0}^{\infty} W^{m^2+m}\right)^5} \\ &= \frac{1}{2^5} e^{\frac{3}{2}\pi i} (1 - 5W^2 + \dots) W^{-\frac{1}{2}} \bar{\Theta}_3'\left(\frac{1}{2} + \frac{1}{2} T | T\right) \end{aligned}$$

und

$$\begin{aligned} \frac{1}{2\pi i} \bar{\Theta}_3'\left(\frac{1}{2} + \frac{1}{2} T | T\right) &= 2 \sum_{\substack{m=-\infty \\ m \equiv 0(3)}}^{+\infty} (-1)^m m W^{\frac{m(m+3)}{3}} + \\ &+ 6 \sum_{\substack{n=1 \\ 2n+m^2 \equiv 0(3)}}^{\infty} \sum_{m=-\infty}^{+\infty} (-1)^m m \psi_3(n) W^{\frac{2n+m(m+3)}{3}} \end{aligned}$$

Denn in dieser letzten Entwicklung kommen keine negativen Potenzen von W vor, während man das von W unabhängige Glied bekommt, wenn man aus der ersten Summe die Glieder $m = 0$ und $m = -3$ nimmt und aus der zweiten Summe die beiden Glieder $m = -1, n = 1$ und $m = -2, n = 1$. Dieses von W unabhängige Glied ist also

$$2 \cdot 3 + 6 [- (-1) \psi_3(1) - 2 \psi_3(1)] = 6 + 6(1 - 2) = 0.$$

Überdies kommen nur ganze Potenzen von W vor, so daß

$$\lim_{T \rightarrow i\infty} W^{-\frac{1}{2}} \bar{\Theta}_3'\left(\frac{1}{2} + \frac{1}{2} T | T\right) = 0$$

und damit auch

$$\lim_{T \rightarrow i\infty} \gamma_3\left(1 - \frac{1}{T}\right) = 0$$

ist.

Jetzt ergibt sich, daß die elliptische Funktion

$$\frac{\bar{\Theta}_3(v|\tau)}{\bar{\theta}^3(v|\tau)}$$

in ihrem Periodenparallelogramm höchstens einen einfachen Pol besitzt und deshalb konstant sein muß, so daß

$$\frac{\bar{\Theta}_3(v|\tau)}{\bar{\theta}^3(v|\tau)} = \frac{\bar{\Theta}_3(0|\tau)}{\bar{\theta}^3(0|\tau)}$$

ist. Schließlich betrachte man die zur Gruppe Γ_3 gehörige Modulfunktion

$$\Phi_3(\tau) = \frac{\bar{\Theta}_3(0|\tau)}{\bar{\theta}^3(0|\tau)}.$$

Aus (92) folgt:

$$\lim_{\tau \rightarrow i\infty} \bar{\Theta}_3(0|\tau) = 2,$$

und weil

$$\lim_{\tau \rightarrow i\infty} \bar{\theta}(0|\tau) = 1$$

ist, ist also

$$\lim_{\tau \rightarrow i\infty} \Phi_3(\tau) = 2.$$

Weiter folgt aus

$$\Phi_3\left(0 \middle| 1 - \frac{1}{T}\right) = \frac{\bar{\Theta}_3\left(\frac{1}{2}T|T\right)}{\bar{\theta}^3\left(\frac{1}{2}T|T\right)}$$

und

$$\bar{\Theta}_3\left(\frac{1}{2}T|T\right) = 2 \sum_{\substack{m=-\infty \\ m \equiv 0(3)}}^{+\infty} W^{\frac{m(m+3)}{3}} + 6 \sum_{\substack{n=1 \\ 2n+m^2 \equiv 0(3)}}^{\infty} \sum_{m=-\infty}^{+\infty} \psi_3(n) W^{\frac{2n+m(m+3)}{3}}$$

leicht, daß auch

$$\lim_{T \rightarrow i\infty} \Phi_3\left(1 - \frac{1}{T}\right) = 2$$

ist. Es ist deshalb $\Phi_3(\tau)$ identisch gleich 2, womit Satz 11 bewiesen ist.

Aus Satz 11 folgt:

Hauptsatz III. Es sei $r_3(n, m)$ die Lösungszahl des diophantischen Gleichungssystems

$$x_1^2 + x_2^2 + x_3^2 = n, \quad x_1 + x_2 + x_3 = m,$$

wo m und n gegebene ganze Zahlen sind. Dann ist $r_3(n, m)$ nur dann $\neq 0$, wenn $m^2 \leq 3n$ und $m \equiv n \pmod{2}$ ist. Falls $m^2 = 3n$ ist, ist $r_3(n, m) = 1$. Falls aber $m^2 < 3n$ und $m \equiv n \pmod{2}$ ist, so ist

$$r_3(n, m) = \alpha_N \sum_{d|N} \left(\frac{d}{3}\right),$$

wo $N = \frac{1}{2}(3n - m^2)$ und $\alpha_N = 3$ oder $= 6$ ist, je nachdem $N \not\equiv 0 \pmod{3}$ oder $N \equiv 0 \pmod{3}$ ist.

Beweis. Die ersten beiden Behauptungen sind trivial bzw. schon in § 1 bewiesen. Es sei also $m^2 < 3n$ und $m \equiv n \pmod{2}$. Aus Satz 11 und (92) folgt dann durch Koeffizientenvergleich

$$r_3(n, m) = 3 \psi_3(N) = 3 \psi_3(-2N; N) = 3 \sum_{\substack{d|N \\ d \not\equiv 0 \pmod{3}}} \left(\frac{d}{3}\right) + 3 \sum_{\substack{d|N \\ d \equiv 0 \pmod{3}}} \left(\frac{N/d}{3}\right).$$

Falls $N \not\equiv 0 \pmod{3}$ ist, ist also

$$r_3(n, m) = 3 \sum_{d|N} \left(\frac{d}{3}\right).$$

Falls N durch 3 teilbar ist, sei 3^β die höchste, in N aufgehende Potenz von 3 und $N = 3^\beta N_0$. In der Summe

$$\sum_{\substack{d|N \\ d \equiv 0 \pmod{3}}} \left(\frac{N/d}{3}\right)$$

sind dann nur diejenigen Glieder $\neq 0$, für die $d = 3^\beta d_0$ und $d_0 | N_0$ ist. Es ist diese Summe also gleich

$$\sum_{d|N_0} \left(\frac{N_0/d}{3}\right).$$

Weil aber N_0/d mit d sämtliche Teiler von N_0 durchläuft, ist sie auch gleich

$$\sum_{d|N_0} \left(\frac{d}{3}\right).$$

Deshalb wird jetzt

$$r_3(n, m) = 6 \sum_{d|N_0} \left(\frac{d}{3}\right).$$

Hierfür kann man aber auch schreiben

$$r_3(n, m) = 6 \sum_{d|N} \left(\frac{d}{3}\right).$$

Denn in der letzteren Summe verschwinden doch diejenigen Glieder, für die d nicht in N_0 aufgeht [also für die $d \equiv 0 \pmod{3}$ ist].

§ 10.

Anwendung auf die Darstellung natürlicher Zahlen als Summe von drei, fünf oder sieben Quadraten.

In diesem Paragraphen sei $s = 3, 5$ oder 7 . Es sei $r_s(n)$ die Anzahl der Darstellungen der natürlichen Zahl n als einer Summe von s Quadraten. Dann ist

$$r_s(n) = \sum_{-\sqrt{s}n \leq m \leq +\sqrt{s}n} r_s(m, n).$$

Dadurch ergeben sich neue Formeln für die Anzahl der Darstellungen einer natürlichen Zahl als einer Summe von 3, 5 oder 7 Quadraten. Aus den Hauptsätzen der vorigen Paragraphen ergeben sich sofort folgende Resultate:

1. Die Anzahl der Darstellungen der natürlichen Zahl n als einer Summe von drei Quadraten ist

$$6 \sum'_{\substack{0 \leq m < \sqrt{3n} \\ m \equiv n \pmod{2}}} \psi_3 \left(\frac{3n - m^2}{2} \right) + 2\lambda_n.$$

Hier bedeutet der Akzent, daß das eventuell (für gerades n) vorkommende Glied $m = 0$ mit dem Faktor $\frac{1}{2}$ zu nehmen ist. Weiter ist $\lambda_n = 1$, falls n durch 3 teilbar und überdies $\frac{1}{3}n$ ein Quadrat ist, und sonst ist $\lambda_n = 0$. Die Funktion $\psi_3(N)$ wird durch

$$\psi_3(N) = \sum_{\substack{d|N \\ d \not\equiv 0 \pmod{3}}} \left(\frac{d}{3} \right) + \sum_{d \equiv 0 \pmod{3}} \left(\frac{N/d}{3} \right)$$

definiert.

2. Die Anzahl der Darstellungen der natürlichen Zahl n als einer Summe von fünf Quadraten ist

$$5 \sum'_{\substack{0 \leq m < \sqrt{5n} \\ m \equiv n \pmod{2}}} (5n - m^2) \psi_5 \left(\frac{5n - m^2}{2} \right) + 2\lambda_n.$$

Hier bedeutet der Akzent, daß das eventuell (für gerades n) vorkommende Glied $m = 0$ mit dem Faktor $\frac{1}{2}$ zu nehmen ist. Weiter ist $\lambda_n = 1$, falls n durch 5 teilbar und überdies $\frac{1}{5}n$ ein Quadrat ist, und sonst ist $\lambda_n = 0$. Die Funktion $\psi_5(N)$ wird durch

$$\psi_5(N) = \sum_{\substack{d|N \\ d \not\equiv 0 \pmod{5}}} \frac{1}{d} \left(\frac{d}{5} \right) - \sum_{d \equiv 0 \pmod{5}} \frac{1}{d} \left(\frac{N/d}{5} \right)$$

definiert.

3. Die Anzahl der Darstellungen der natürlichen Zahl n als einer Summe von sieben Quadraten ist

$$\frac{7}{16} \sum'_{\substack{0 \leq m < \sqrt{7n} \\ m \equiv n \pmod{2}}} (7n - m^2)^2 \psi_7 \left(\frac{7n - m^2}{2} \right) + 2\lambda_n.$$

Hier bedeutet der Akzent, daß das eventuell (für gerades n) vorkommende Glied $m = 0$ mit dem Faktor $\frac{1}{2}$ zu nehmen ist. Weiter ist $\lambda_n = 1$, falls n durch 7 teilbar und überdies $\frac{1}{7}n$ ein Quadrat ist, und sonst ist $\lambda_n = 0$. Die Funktion $\psi_7(N)$ wird durch

$$\psi_7(N) = \sum_{\substack{d|N \\ d \not\equiv 0 \pmod{7}}} \frac{1}{d^3} \left(\frac{d}{7} \right) - \sum_{d \equiv 0 \pmod{7}} \frac{1}{d^3} \left(\frac{N/d}{7} \right)$$

definiert.

Über die dreigliedrigen integrierbaren Gruppen.

Von

B. von Kerékjártó in Budapest.

Eine n -gliedrige kontinuierliche Gruppe G_n heißt *integrierbar*, wenn sie eine $n - 1$ -gliedrige invariante Untergruppe G_{n-1} , dieselbe eine $n - 2$ -gliedrige invariante Untergruppe G_{n-2} besitzt, usf.; laut Definition ist jede eingliedrige kontinuierliche Gruppe integrierbar. Aus der Aufzählung der zweigliedrigen kontinuierlichen Gruppen ergibt sich, daß auch jede zweigliedrige kontinuierliche Gruppe G_2 integrierbar ist. Wenn nämlich G_2 kommutativ ist, so ist jede eingliedrige Untergruppe invariant in G_2 ; wenn aber G_2 nicht-kommutativ ist, so ist sie einstufig und stetig isomorph der Ähnlichkeitsgruppe der Geraden, in der die Verschiebungen der Geraden eine eingliedrige invariante Untergruppe bilden¹⁾.

Zweck der vorliegenden Arbeit ist, die verschiedenen Typen der dreigliedrigen integrierbaren Gruppen zu bestimmen mittels der expliziten Darstellungen der zur Gruppe gehörigen Transformationen. Die einfachsten bekannten Gruppen dieser Art sind die Gruppe der Verschiebungen des dreidimensionalen euklidischen Raumes, ferner die Gruppe der Bewegungen und die Gruppe der homothetischen Abbildungen der euklidischen Ebene auf sich selbst; die beiden letzteren enthalten die Gruppe der Verschiebungen der euklidischen Ebene als zweigliedrige invariante Untergruppe.

Bekanntlich spielen die integrierbaren Gruppen in der Theorie der Differentialgleichungen eine ähnliche wichtige Rolle, wie die Galoisschen auflösbaren Gruppen in der Theorie der algebraischen Gleichungen²⁾. Auf die Anwendungen der unten mitgeteilten Resultate werde ich bei einer anderen Gelegenheit zurückkommen.

§ 1.

Die in einer dreigliedrigen integrierbaren Gruppe G_3 enthaltenen zweigliedrigen Untergruppen.

Sei G_3 eine dreigliedrige integrierbare Gruppe und sei G_2 eine zweigliedrige invariante Untergruppe von G_3 . Die Faktorgruppe G_3/G_2 ist eine eingliedrige kontinuierliche Gruppe und also homöomorph entweder der Gruppe der

¹⁾ B. von Kerékjártó, Geometrische Theorie der zweigliedrigen kontinuierlichen Gruppen. Abhandl. Math. Sem. Hamburg 8 (1930), S. 107–114.

²⁾ E. Picard, Traité d'Analyse, III. (Paris 1896). Kapitel XVII.

Verschiebungen einer Geraden, oder der Gruppe der Drehungen eines Kreises. Bezeichnen wir mit γ den (bis auf einen konstanten Faktor bestimmten) kanonischen Parameter der Gruppe G_3/G_2 .

Die Gruppe G_2 enthält eine eingliedrige invariante Untergruppe $G_1 = (S^a)$, die aus sämtlichen Potenzen S^a ($a = \text{reell}$) eines Elementes S besteht. Sei T ein nicht zu (S^a) gehöriges Element von G_2 , und sei (T^β) die von ihm erzeugte eingliedrige kontinuierliche Gruppe. Die Gruppe G_2 ist das Produkt von (T^β) und (S^a) : Jedes Element von G_2 ist auf eine und nur auf eine Weise in der Form $T^\beta S^a$ darstellbar. Aus der Invarianz der Untergruppe (S^a) in G_2 folgt ferner, daß für jedes reelle Zahlenpaar α, β die Beziehung gilt:

$$T^{-\beta} S^a T^\beta = S^{a^\beta}, \text{ also } S^a T^\beta = T^\beta S^{a^\beta} \text{ und } T^\beta S^a = S^{a a^{-\beta}} T^\beta,$$

wo a die durch die Gleichung

$$S^a = T^{-1} S T$$

bestimmte positive reelle Zahl bedeutet; es ist $a = 1$ bzw. $a \neq 1$, je nachdem G_2 kommutativ ist oder nicht¹⁾.

Wenn G_2 nicht-kommutativ ist, so ist sie der Parametergruppe der Ähnlichkeitsgruppe der Geraden homöomorph; in diesem Fall ist der Wertvorrat der Parameter: $-\infty < \alpha < +\infty, -\infty < \beta < +\infty$. Wenn G_2 kommutativ ist, so ist sie homöomorph der Gruppe der Verschiebungen der euklidischen Ebene, oder des Kreiszylinders, oder des Torus; den drei Fällen entsprechen die folgenden Wertvorräte der Parameter: 1. $-\infty < \alpha < +\infty, -\infty < \beta < +\infty$; 2. $0 \leq \alpha < 1, -\infty < \beta < +\infty$; 3. $0 \leq \alpha < 1, 0 \leq \beta < 1$.

Nehmen wir an, daß die zweigliedrige invariante Untergruppe G_2 von G_3 nicht-kommutativ ist. Wir werden nachweisen, daß G_3 auch eine kommutative invariante zweigliedrige Untergruppe enthält.

1. 1. Wenn ein Element A von G_2 mit den Elementen der eingliedrigen invarianten Untergruppe (S^a) von G_2 vertauschbar ist, so gehört A zu (S^a) .

Sei $A = T^\beta S^a$; wenn A mit $S^{a'}$ vertauschbar ist, so ist:

$$(T^\beta S^a)^{-1} S^{a'} (T^\beta S^a) = S^{-a} (T^{-\beta} S^{a'} T^\beta) S^a = S^{a' a^\beta} = S^{a'},$$

und weil $a \neq 1$, so ist $\beta = 0$, d. h. $A = S^a$.

Der Kommutator $A^{-1} B^{-1} A B$ von je zwei Elementen A und B von G_2 gehört zu (S^a) ; sei nämlich $A = T^\beta S^a$ und $B = T^{\beta'} S^{a'}$; ihr Kommutator:

$$(T^\beta S^a)^{-1} (T^{\beta'} S^{a'})^{-1} (T^\beta S^a) (T^{\beta'} S^{a'}) = S^{-a} T^{-\beta} S^{-a'} T^{-\beta'} T^\beta S^a T^{\beta'} S^{a'}$$

läßt sich infolge der Vertauschbarkeit von T^β und $T^{\beta'}$ mit der Gruppe (S^a)

auf die Gestalt S^a bringen. Andererseits ist ein beliebiges Element S^a der Gruppe (S^a) der Kommutator der Elemente T^j und S^a von G_2 , wenn nämlich

$$\alpha > |\alpha'| \quad \text{und} \quad \beta = \frac{\log(\alpha - \alpha') - \log \alpha}{\log a}$$

gesetzt wird. (S^a) ist also die Kommutatorgruppe von G_2 .

Wenn R ein beliebiges Element von G_3 ist, geht die Kommutatorgruppe von G_2 bei Transformation durch R in sich selbst über. Dem Kommutator der Elemente A und B von G_2 entspricht nämlich der Kommutator der ebenfalls zu G_2 gehörigen Elemente $A_1 = R^{-1}AR$ und $B_1 = R^{-1}BR$, laut der Formel:

$$\begin{aligned} R^{-1}(A^{-1}B^{-1}AB)R &= (R^{-1}A^{-1}R)(R^{-1}B^{-1}R)(R^{-1}AR)(R^{-1}BR) \\ &= A_1^{-1}B_1^{-1}A_1B_1. \end{aligned}$$

Daraus folgt, daß die eingliedrige Gruppe $G_1 = (S^a)$ nicht nur in G_2 , sondern auch in G_3 eine invariante Untergruppe ist.

1. 2. Wenn ein Element R von G_3 mit S vertauschbar ist, dann ist es auch mit jedem Element der Gruppe (S^a) vertauschbar.

Die Vertauschbarkeit von R und S drückt sich nämlich durch die Gleichung: $R^{-1}SR = S$ aus; durch Potenzieren folgt daraus: $R^{-1}S^2R = S^2$, $R^{-1}S^3R = S^3$, ..., $R^{-1}S^{-1}R = S^{-1}$, ... Ferner ergibt sich: $(R^{-1}S^{1/2}R)^2 = R^{-1}SR = S$, also $R^{-1}S^{1/2}R = S^{1/2}$, und ebenso: $R^{-1}S^{1/2^k}R = S^{1/2^k}$, $R^{-1}S^{a/2^k}R = S^{a/2^k}$; da $R^{-1}S^aR$ sich stetig mit S^a ändert, folgt daraus für jede reelle Zahl α : $R^{-1}S^aR = S^a$.

1. 3. Diejenigen Elemente von G_3 , die mit den Elementen von (S^a) vertauschbar sind, bilden eine Gruppe Γ .

Wenn nämlich

$$R^{-1}SR = S \quad \text{und} \quad R_1^{-1}SR_1 = S,$$

so folgt

$$(RR_1)^{-1}S(RR_1) = S.$$

Zwecks anschaulicheren Ausdrucks werden wir im Gruppenraum von G_3 die Menge der zur Gruppe G_2 bzw. zur Nebengruppe RG_2 gehörigen Elemente als Ebene G_2 bzw. RG_2 , ferner die Menge der zur eingliedrigen Untergruppe $G_1 = (S^a)$ von G_2 bzw. zur Nebengruppe RG_1 gehörigen Elemente als Gerade G_1 bzw. RG_1 bezeichnen.

1. 4. Wenn R zur Gruppe Γ gehört, so gehört jeder Punkt RS^a der Geraden RG_1 zu Γ . Die Ebene RG_2 enthält außer der Geraden RG_1 kein weiteres Element der Gruppe Γ .

Wenn nämlich R zu Γ gehört, so ist $R^{-1}SR = S$ und

$$(RS^a)^{-1}S(RS^a) = S^{-a}(R^{-1}SR)S^a = S^{-a}SS^a = S,$$

also gehört das Element RS^a ebenfalls zu Γ . — Wenn R und $RT^\beta S^a$ zu Γ gehören, so ist

$$\begin{aligned} S &= (RT^\beta S^a)^{-1}S(RT^\beta S^a) = S^{-a}T^{-\beta}(R^{-1}SR)T^\beta S^a = S^{-a}(T^{-\beta}ST^\beta)S^a \\ &= S^{-a}S^{a^\beta}S^a = S^{a^\beta}, \end{aligned}$$

also $a^\beta = 1$, und da $a \neq 1$, so ist $\beta = 0$.

1. 5. Jede Ebene RG_2 enthält eine zu Γ gehörige Gerade $R'G_1$; dieselbe ändert sich stetig mit der Ebene RG_2 .

Sei R ein beliebiges Element von G_3 ; zufolge der Invarianz der Untergruppe (S^a) in G_3 besteht die Beziehung:

$$R^{-1}SR = S^{a'}.$$

Wenn

$$\beta = -\frac{\log a'}{\log a}$$

angenommen wird, so ist

$$T^{-\beta}S^{a'}T^\beta = S^{a'a^\beta} = S,$$

und

$$(RT^\beta)^{-1}S(RT^\beta) = T^{-\beta}(R^{-1}SR)T^\beta = T^{-\beta}S^{a'}T^\beta = S;$$

$R' = RT^\beta$ gehört somit zu Γ , und also auch die Gerade $R'G_1$.

Aus der Stetigkeit der Gruppe G_3 folgt, daß für eine beliebige gegen R' konvergierende Folge von Elementen $R^{(1)}, R^{(2)}, \dots$ der Gruppe G_3 die Geraden $R^{(1)}G_1, R^{(2)}G_1, \dots$ gleichmäßig gegen die Gerade $R'G_1$ konvergieren, d. h. daß der Parameterabstand der Geraden $R^{(k)}G_1$ und $R'G_1$ (in bezug auf eine beschränkte Metrik im Gruppenraum) mit wachsendem k gegen 0 konvergiert. Um die stetige Änderung der zu RG_2 gehörigen Geraden $R'G_1$ von Γ mit der Ebene RG_2 zu beweisen, genügt es zu zeigen, daß für jedes Element R von G_3 , das genügend wenig von R entfernt ist, die zur Ebene RG_2 gehörige Gerade $\bar{R}'G_1$ von Γ einen Punkt in beliebiger Nähe der Geraden $R'G_1$ besitzt.

Für das im Gruppenraum veränderliche Element X erklären wir die eindeutige stetige Funktion $f(X)$ mittels der Formel:

$$X^{-1}SX = S^{f(X)}.$$

Wenn X zu Γ gehört, und nur dann, ist $f(X) = 1$. Die Ebene RG_2 wird durch ihre zu Γ gehörige Gerade $R'G_1$ in zwei Halbebenen zerlegt; in einer derselben ist $f(X) > 1$, in der anderen $f(X) < 1$. Sei U eine (beliebig kleine) räumliche Umgebung von R' ; seien L_1 und L_2 zwei solche Punkte von U , die zur Ebene RG_2 gehören, in derselben durch die Gerade $R'G_1$ getrennt sind und durch einen in U und zugleich in RG_2 liegenden stetigen Bogen mit einander ver-

bunden werden können. Sei V_1 und V_2 je eine räumliche Umgebung von L_1 und von L_2 , Teilbereiche von U , in denen $f(X) > 1$ bzw. $f(X) < 1$ (sowie in L_1 bzw. L_2). Wenn \bar{R} genügend wenig von R entfernt ist, gibt es in der Ebene $\bar{R}G_2$ je einen zu V_1 bzw. zu V_2 gehörigen Punkt und einen die beiden Punkte in U verbindenden stetigen Bogen. Dieser Bogen besitzt infolge der Stetigkeit von $f(X)$ mindestens einen Punkt X , in dem $f(X) = 1$ ist, d. h. einen Punkt der zur Ebene $\bar{R}G_2$ gehörigen Geraden von Γ .

Die zu Γ gehörigen Geraden $R'G_1$ hängen also vom Parameter γ der Faktorgruppe G_3/G_2 stetig ab und bilden eine *reguläre Kurvenschar*. Nach einer Methode, die ich an anderer Stelle entwickelt habe³⁾, läßt sich also ein einfacher stetiger Bogen λ in Γ derart angeben, daß er mit jeder zu Γ gehörigen Geraden $R'G_1$ je ein und nur ein Element gemeinsam hat.

Ordnen wir jedem Punkte R' von λ den der Ebene $R'G_2$ entsprechenden Parameterwert γ von G_3/G_2 , und jedem Punkt $R'S^\alpha$ von Γ die Parameter (α, γ) zu. Das Entsprechen der Zahlenpaare (α, γ) und der Elemente der Gruppe Γ ist topologisch. Daraus folgt, daß Γ eine *zweigliedrige kontinuierliche Gruppe* ist. Weil jedes Element von Γ mit den Elementen der eingliedrigen Untergruppe (S^α) von Γ vertauschbar ist, folgt aus 1. 1, daß Γ eine *kommutative Gruppe* ist.

Γ ist eine invariante Untergruppe von G_3 . Wenn nämlich L zu G_3 , und R zu Γ gehört, so ist $R^{-1}SR = S$, laut der Erklärung von Γ , und $L^{-1}SL = S^\alpha$, infolge der Invarianz von (S^α) in G_3 . Daraus folgt:

$$\begin{aligned} S^{\alpha'} &= L^{-1}SL = L^{-1}(R^{-1}SR)L \\ &= (L^{-1}R^{-1}L)(L^{-1}SL)(L^{-1}RL) = R_1^{-1}S^{\alpha'}R_1, \end{aligned}$$

wo $R_1 = L^{-1}RL$ das Transformierte von R durch L bedeutet; R_1 ist vertauschbar mit $S^{\alpha'}$, gehört also zu Γ .

Somit haben wir den folgenden Satz bewiesen:

Satz 1. *Jede dreigliedrige integrierbare Gruppe enthält eine kommutative zweigliedrige invariante Untergruppe.*

§ 2.

Über die eingliedrigen Untergruppen von G_3 .

Laut des Satzes 1 nehmen wir an, daß G_2 eine *kommutative zweigliedrige invariante Untergruppe* von G_3 ist. Setzen wir voraus, daß G_2 der Gruppe der Verschiebungen der Ebene homöomorph ist; die beiden anderen Möglichkeiten, denen entsprechend der Gruppenraum von G_2 einer Zylinderfläche oder einer Torusfläche homöomorph ist, lassen sich leicht auf diese zurückführen.

³⁾ B. von Kerékjártó, Vorlesungen über Topologie, I (Berlin, 1923), S. 242–245.

Sei R ein beliebiges Element von G_3 ; dem veränderlichen Punkt X des Gruppenraumes lassen wir den Bildpunkt $X' = R^{-1}XR$ entsprechen; dadurch erklären wir eine topologische Abbildung \mathfrak{A}_R des Gruppenraumes auf sich selbst. Die den verschiedenen Elementen R, R', \dots von G_3 entsprechenden Abbildungen $\mathfrak{A}_R, \mathfrak{A}_{R'}, \dots$ bilden die *adjungierte Gruppe* von G_3 . Wenn G_3 kommutativ ist, besteht die adjungierte Gruppe aus der Identität als einzigem Element.

Bei jeder Abbildung der adjungierten Gruppe geht die Ebene G_2 in sich über. Wenn S ein beliebiges Element von G_2 und X einen veränderlichen Punkt der Ebene G_2 bedeutet, so besteht für jedes Element R von G_3 :

$$(RS)^{-1}X(RS) = S^{-1}(R^{-1}XR)S = S^{-1}X'S = X' = R^{-1}XR;$$

das bedeutet, daß zwei beliebige Elemente R und RS von G_3 , die zur selben Nebengruppe RG_2 gehören, eine und dieselbe Abbildung der adjungierten Gruppe in der Ebene G_2 erzeugen; diese Abbildung bezeichnen wir mit \mathfrak{A}_γ , wo γ den der Ebene RG_2 entsprechenden Parameterwert bedeutet.

Daraus folgt, daß, falls G_3 nicht-kommutativ ist, die adjungierte Gruppe eine eingliedrige kontinuierliche Gruppe von topologischen Abbildungen der Ebene G_2 in sich erzeugt.

Der Zweck der weiteren Überlegungen ist nachzuweisen, daß G_3 mindestens eine nicht zu G_2 gehörige eingliedrige kontinuierliche Untergruppe enthält.

Wir erklären eine eindeutige und stetige Abbildung Ω des Gruppenraumes in sich, indem wir jedem Element X sein Quadrat $X' = X^2$ als Bild zuordnen. Bei dieser Abbildung geht die Ebene G_2 in sich, und die Ebene RG_2 , welche dem Parameter γ entspricht, in die Ebene R^2G_2 über, welche dem Parameter 2γ entspricht; das folgt unmittelbar aus der Annahme, daß γ der kanonische Parameter der Faktorgruppe G_3/G_2 ist. — Wir werden die Ebene RG_2 , die dem Parameterwert γ entspricht, auch als *Ebene γ* bezeichnen.

Wir behaupten, daß die durch Ω erzeugte eindeutige und stetige Abbildung der Ebene γ auf die Ebene 2γ umkehrbar eindeutig ist, wenn γ hinreichend klein ist. Unter der entgegengesetzten Annahme seien R und RS zwei solche Elemente der Nebengruppe RG_2 , für welche

$$R^2 = (RS)^2.$$

Daraus folgt:

$$R^{-1}SR = S^{-1}.$$

Laut des Beweises vom Hilfssatz 1.2 folgt daraus für jede reelle Zahl α :

$$R^{-1}S^\alpha R = S^{-\alpha};$$

das bedeutet, daß die Abbildung \mathfrak{A}_R eine Spiegelung der Geraden (S^α) in bezug auf den Punkt I (Identität) erzeugt. Falls aber der Parameter γ der

Nebengruppe RG_2 hinreichend klein ist, unterscheidet sich \mathfrak{A}_n beliebig wenig von der identischen Abbildung, kann also nicht eine Spiegelung der Geraden (S^α) erzeugen.

Wir nehmen eine solche Zahl γ_0 , daß die Abbildung Ω den zwischen den Ebenen γ_0 und $-\gamma_0$ liegenden (den Punkt I enthaltenden) Bereich des Gruppenraumes auf den zwischen den Ebenen $2\gamma_0$ und $-2\gamma_0$ liegenden Bereich, oder auf einen Teilbereich desselben topologisch abbildet.

Wir bestimmen eine Umgebung U der Identität I derart, daß U im Innern ihres bei Ω entstehenden Bildes $\Omega(U)$ liegt. Dazu nehmen wir eine solche einfache geschlossene Kurve c in der Ebene G_2 an, die im Innern der Bildkurve $c' = \Omega(c)$ liegt. Wenn (S^α) und (T^β) zwei eingliedrige Untergruppen von G_2 sind, so bilden die Punkte $T^\beta S^{\pm\alpha_0}$ ($-\beta_0 \leq \beta \leq \beta_0$) und $T^{\pm\beta_0} S^\alpha$ ($-\alpha_0 \leq \alpha \leq \alpha_0$) eine solche Kurve c , deren Bild c' aus den Punkten $T^\beta S^{\pm 2\alpha_0}$ ($-2\beta_0 \leq \beta \leq 2\beta_0$) und $T^{\pm 2\beta_0} S^\alpha$ ($-2\alpha_0 \leq \alpha \leq 2\alpha_0$) besteht; c liegt also im Innern von $c' = \Omega(c)$. Da die abgeschlossenen Mengen c und c' keinen gemeinsamen Punkt haben, können wir eine räumliche Umgebung W von c derart annehmen, daß W und $W' = \Omega(W)$ keinen gemeinsamen Punkt haben. — Sei V eine hinreichend kleine Umgebung der Identität, so daß das Produkt $RT^\beta S^\alpha$ eines beliebigen Elementes R von V mit einem beliebigen zu c gehörigen Element $T^\beta S^\alpha$ zu W gehört. — Wir verbinden in V zwei Punkte, die in den Ebenen γ_1 und $-\gamma_1$ liegen, durch einen Bogen λ ; dabei setzen wir voraus, daß $0 < \gamma_1 < \gamma_0/2$ ist. Sei R ein veränderlicher Punkt von λ und $T^\beta S^\alpha$ ein veränderlicher Punkt von c ; wir bilden sämtliche Produkte $RT^\beta S^\alpha$; die entsprechenden Punkte bilden eine in W liegende „Zylinderfläche“, die samt je einem Bereich der Ebenen γ_1 und $-\gamma_1$ eine Umgebung U der Identität I abgrenzt. Das Bild des Randes von U bei der Abbildung Ω besteht aus einer in W' liegenden „Zylinderfläche“ und je einem in der Ebene $2\gamma_1$ bzw. $-2\gamma_1$ liegenden Bereich; der Rand von U liegt im Innern seines bei Ω entstehenden Bildes. Weil die Abbildung Ω im Zwischengebiet der Ebenen γ_1 und $-\gamma_1$ topologisch ist, so folgt daraus, daß die Umgebung U bei Ω topologisch in einen U enthaltenden Bereich $\Omega(U)$ übergeht.

Bei den negativen Potenzen von Ω geht also U in solche Bereiche

$$U^{-1} = \Omega^{-1}(U), \quad U^{-2} = \Omega^{-2}(U), \dots$$

über, von denen jeder im vorangehenden enthalten ist.

I ist der einzige gemeinsame Punkt sämtlicher Bereiche U^{-n} .

Wäre nämlich $R \neq I$ ein anderer gemeinsamer Punkt sämtlicher Bereiche U^{-n} , so würde $\Omega^n(R) = R^{2^n}$ für jedes n zu U gehören. Wenn aber R zur Ebene G_2 gehört, so ist die Folge R^{2^n} divergent¹⁾. Wenn dagegen R nicht zur Ebene G_2 gehört, sei γ der Parameterwert der Nebengruppe RG_2 , und es

sei n die kleinste ganze Zahl, für welche $2^n |\gamma| > \gamma_1$; es ist dann $\gamma_1 < 2^n |\gamma| < \gamma_0$ und R^{2^n} kann also nicht zu U gehören.

2. 1. Wenn also R ein beliebiges Element von U ist, so konvergiert die Folge

$$R^{1/2} = Q^{-1}(R), \quad R^{1/4} = Q^{-2}(R), \dots$$

gegen die Identität I .

Bezeichnen wir mit $R^{n/2^k}$ die n -te Potenz von $R^{1/2^k}$ ($n = 0, \pm 1, \pm 2, \dots$).

2. 2. Sei $\gamma_1 = n_1/2^{k_1}$, $\gamma_2 = n_2/2^{k_2}$, ... eine gegen 0 konvergierende Folge von dyadisch rationalen Zahlen. Wenn die Folge $R^{\gamma_1}, R^{\gamma_2}, \dots$ nicht gegen I konvergiert, so läßt sich eine gegen 0 konvergierende Folge von dyadisch rationalen Zahlen $\gamma'_1, \gamma'_2, \dots$ derart angeben, daß die Folge $R^{\gamma'_1}, R^{\gamma'_2}, \dots$ gegen ein von I verschiedenes Element der Ebene G_2 konvergiert.

Wenn die Folge $R^{\gamma_1}, R^{\gamma_2}, \dots$ ein von I verschiedenes Häufungselement besitzt, so gehört es notwendig zur Ebene G_2 ; dann ist die obige Behauptung evident. Nehmen wir also an, daß die Folge $R^{\gamma_1}, R^{\gamma_2}, \dots$ divergent ist. Sei U eine kompakte Umgebung von I ; bezeichnen wir mit U^2 die Menge derjenigen Elemente, die als Produkt von zwei zu U gehörigen Elementen darstellbar sind. Da die Folge $R^{1/2^v}$ ($v = 1, 2, \dots$) gegen I konvergiert, können wir annehmen, daß sämtliche $R^{1/2^v}$ zu U gehören. Bedenke für jedes v, m , die kleinste positive ganze Zahl, für welche $R^{m/2^k}$ nicht zu U gehört. Weil $R^{(m-1)/2^k}$ und $R^{1/2^k}$ zu U gehören, ist $R^{m/2^k}$ in U^2 enthalten. Die Menge U^2 ist kompakt, folglich besitzt die Menge $R^{m/2^k}$ ($v = 1, 2, \dots$) mindestens ein Häufungselement, und da die Zahlfolge $m_v/2^{k_v}$ gegen 0 konvergiert, gehört ein solches notwendig zur Ebene G_2 . Da aber kein Element der Folge $R^{m_v/2^{k_v}}$ in U liegt, ist jedes Häufungselement der Folge von I verschieden. Daraus folgt die Behauptung 2. 2.

Bezeichnen wir mit Σ die Menge derjenigen Elemente von G_2 , die mit R vertauschbar sind. Die Elemente von Σ und nur diese sind invariant bei der topologischen Abbildung des Gruppenraumes in sich, erklärt durch die Formel: $X' = R^{-1}XR$. Daraus folgt, daß Σ eine abgeschlossene Menge ist. Es ist klar, daß die Wurzel $R^{1/2^k}$ und die Potenzen $R^{n/2^k}$ mit R vertauschbar sind, also zu Σ gehören; auch jedes Häufungselement der Menge der $R^{n/2^k}$ gehört also zu Σ .

Betreffend die Ausdehnung von Σ sind die folgenden drei Fälle möglich.

1. Σ hat in der Ebene G_2 kein Element außer I .

In diesem Falle konvergiert, für jede gegen 0 konvergierende Folge dyadisch rationaler Zahlen $\gamma_1, \gamma_2, \dots$, die Folge $R^{\gamma_1}, R^{\gamma_2}, \dots$ gegen I , laut

des Hilfssatzes 2. 2. Daraus ergibt sich, daß die den dyadisch rationalen Exponenten γ entsprechenden Potenzen R^γ und deren Häufungselemente eine eingliedrige kontinuierliche Gruppe ergeben.

2. Σ hat in der Ebene G_2 ein von 1 verschiedenes Element S , und jedes zu G_2 gehörige Element von Σ gehört zur Geraden (S^a) .

Da R und S vertauschbar sind, folgt aus 1. 2, daß für jede reelle Zahl α R und S^a vertauschbar sind; sämtliche Punkte der Geraden (S^a) gehören also zu Σ . Wenn γ eine beliebige dyadisch rationale Zahl ist, gehören sämtliche Punkte der Geraden $R^\gamma(S^a)$ ebenfalls zu Σ ; außer diesen Punkten enthält aber die Ebene $R^\gamma G_2$ keinen weiteren Punkt von Σ . — Wenn die Folge dyadisch rationaler Zahlen $\gamma_1, \gamma_2, \dots$ gegen 0 konvergiert, so konvergieren die Geraden $R^{\gamma_1}(S^a), R^{\gamma_2}(S^a), \dots$ gegen die Gerade (S^a) . (Der Beweis dieser Behauptung ist ähnlich dem vom Hilfssatz 2. 2.) Daraus folgt, daß die auf den Geraden $R^\gamma(S^a)$ liegenden Punkte (γ dyadisch rational) und deren Häufungspunkte eine der Ebene lokal homöomorphe Punktmenge bilden; dieselbe ist die Menge Σ . Da die Elemente von Σ eine Gruppe bilden, so ergibt sich, daß Σ eine zweigliedrige kontinuierliche Gruppe ist, in welcher die Potenzen R^γ (γ reell) eine eingliedrige kontinuierliche Untergruppe bilden.

3. Σ besitzt in der Ebene G_2 zwei solche Elemente S und T , von denen T nicht zur Geraden (S^a) gehört.

Sämtliche Punkte der Geraden (S^a) und (T^b) , ferner sämtliche Punkte $T^b S^a$, d. h. alle Punkte der Ebene G_2 , gehören in diesem Falle zu Σ . Für jede dyadisch rationale Zahl γ gehören sämtliche Punkte der Ebene $R^\gamma G_2$ zu Σ , und weil diese Punkte im Gruppenraum überall dicht liegen, ist die abgeschlossene Menge Σ mit dem Gruppenraum identisch, d. h. jedes Element von G_2 ist mit R vertauschbar. Daraus folgt nach dem Beweis von Hilfssatz 1. 2, daß jedes Element von G_2 mit jedem R^γ , dessen Exponent γ dyadisch rational ist, vertauschbar ist. Insbesondere ist R^γ mit jedem Element von G_2 und mit jedem $R^{\gamma'}$ vertauschbar, wenn γ' ebenfalls dyadisch rational ist.

Seien A und B zwei solche Elemente von G_2 , die zu den den dyadisch rationalen Zahlen γ und γ' entsprechenden Nebengruppen $R^\gamma G_2$ und $R^{\gamma'} G_2$ gehören. Es ist also

$$A = R^\gamma S \quad \text{und} \quad B = R^{\gamma'} S',$$

wo S und S' Elemente von G_2 bedeuten. Das Transformierte von B durch A :

$$A^{-1} B A = (R^\gamma S)^{-1} (R^{\gamma'} S') (R^\gamma S) = S^{-1} R^{-\gamma} R^{\gamma'} S' R^\gamma S$$

läßt sich zufolge der gegenseitigen Vertauschbarkeit von S, S', R^γ und $R^{\gamma'}$ in der Form schreiben:

$$A^{-1} B A = R^{\gamma'} S' = B;$$

das bedeutet, daß A und B miteinander vertauschbar sind.

Wenn A' und B' zwei beliebige Elemente von G_3 sind, so gibt es in deren Nähe solche Elemente A und B , denen entsprechende Parameter γ und γ' dyadisch rationale Zahlen sind, und also laut der obigen $A^{-1}B^{-1}AB = I$. Wegen der Stetigkeit der Gruppe G_3 folgt daraus, daß der Kommutator von A' und B' ebenfalls I ist, daß also G_3 eine kommutative Gruppe ist.

Nach einem Satz von Pontrjagin ⁴⁾ ist die kommutative Gruppe G_3 homöomorph entweder der Gruppe der Verschiebungen des dreidimensionalen euklidischen Raumes, oder der Faktorgruppe derselben in bezug auf die invariante Untergruppe: $\gamma' = \gamma + n$ ($n = 0, \pm 1, \pm 2, \dots$).

Das Resultat der obigen Überlegung sprechen wir im folgenden Satz aus:

Satz 2. G_3 enthält mindestens eine nicht zu G_2 gehörige eingliedrige kontinuierliche Untergruppe (R'). G_3 ist das Produkt der zweigliedrigen kommutativen Gruppe G_2 und der eingliedrigen Gruppe (R').

Jedes Element von G_3 läßt sich also auf eine und nur auf eine Weise in der Form $R' T^\theta S^\alpha$ darstellen, wenn (S^α) und (T^θ) zwei beliebige eingliedrige Untergruppen von G_2 bedeuten.

§ 3.

Die explizite Darstellung der Transformationen von G_3 .

Bezeichnen wir mit X einen veränderlichen Punkt des Gruppenraumes von G_3 , und mit A ein Element von G_3 . Die Ausdrücke derselben seien:

$$X = R^\alpha T^\nu S^\pi \quad \text{und} \quad A = R' T'^\nu S'^\pi.$$

Bei der Abbildung A der Gruppe G_3 geht der Punkt X in $X' = XA$ über; es sei:

$$X' = R^\alpha T'^\nu S'^\pi.$$

Es ist also:

$$R' T'^\nu S'^\pi = R^\alpha T^\nu S^\pi \cdot R' T'^\nu S'^\pi.$$

Der Ausdruck auf der rechten Seite läßt sich folgendermaßen umändern:

$$R^\alpha R' (R^{-\nu} T^\nu S^\pi R') T'^\nu S'^\pi = R^{\alpha+\nu} (R^{-\nu} T^\nu S^\pi R') T'^\nu S'^\pi.$$

Weil $G_2 = (T^\theta S^\alpha)$ eine invariante Untergruppe von G_3 ist, so folgt:

$$(*) \quad R^{-\nu} (T^\nu S^\pi) R' = T^{g_2(x, y, \gamma)} S^{g_1(x, y, \gamma)},$$

wo $g_1(x, y, \gamma)$ und $g_2(x, y, \gamma)$ eindeutige stetige Funktionen der Veränderlichen x, y und des Parameters γ bedeuten. Da G_2 kommutativ ist, ergibt sich

⁴⁾ L. Pontrjagin, Topological groups. (Princeton 1939) S. 170, Satz 43.

mittels dieser Funktionen der folgende Ausdruck der Abbildungen von G_3 :

$$R^{-\gamma} T^{\gamma} S^{\gamma} = R^{\alpha + \gamma} T^{\gamma_2(x, y, \gamma)} S^{\gamma_1(x, y, \gamma)} T^{\gamma} S^{\alpha} = R^{\alpha + \gamma} T^{\gamma_2(x, y, \gamma) + \beta} S^{\gamma_1(x, y, \gamma) + \alpha},$$

also:

$$x' = g_1(x, y, \gamma) + \alpha, \quad y' = g_2(x, y, \gamma) + \beta, \quad z' = z + \gamma.$$

Wenn die Gruppe G_3 kommutativ ist, so ist $g_1(x, y, \gamma) \equiv x$ und $g_2(x, y, \gamma) \equiv y$, und also ist der Ausdruck der Abbildungen von G_3 der folgende:

$$x' = x + \alpha, \quad y' = y + \beta, \quad z' = z + \gamma.$$

Wenn G_3 nicht-kommutativ ist, ersetzen wir zwecks der Bestimmung der Funktionen $g_i(x, y, \gamma)$ in der Gleichung (*) x und y durch $x_1 + x_2$ und $y_1 + y_2$; so erhalten wir:

$$R^{-\gamma} T^{\gamma_1 + \gamma_2} S^{\gamma_1 + \gamma_2} R^{\gamma} = (R^{-\gamma} T^{\gamma_1} R^{\gamma}) (R^{-\gamma} T^{\gamma_2} R^{\gamma}) (R^{-\gamma} S^{\gamma_1} R^{\gamma}) (R^{-\gamma} S^{\gamma_2} R^{\gamma}),$$

woraus mit Rücksicht auf die Kommutativität von G_2 folgt:

$$(**) \quad g_i(x_1 + x_2, y_1 + y_2, \gamma) = g_i(0, y_1, \gamma) + g_i(0, y_2, \gamma) + g_i(x_1, 0, \gamma) + g_i(x_2, 0, \gamma) \quad (i = 1, 2).$$

Wenn $\gamma = 0$, so ist:

$$g_1(x, y, 0) \equiv x, \quad g_2(x, y, 0) \equiv y.$$

In der Gleichung (**) setzen wir der Reihe nach: $y_1 = y_2 = 0$, dann $x_1 = x_2 = 0$, schließlich: $x_1 = x$, $y_2 = y$, $x_2 = y_1 = 0$, so erhalten wir die folgenden Zusammenhänge:

$$g_i(x_1 + x_2, 0, \gamma) = g_i(x_1, 0, \gamma) + g_i(x_2, 0, \gamma),$$

also:

$$g_i(x, 0, \gamma) = x f_{i1}(\gamma);$$

ähnlicherweise:

$$g_i(0, y_1 + y_2, \gamma) = g_i(0, y_1, \gamma) + g_i(0, y_2, \gamma),$$

also:

$$g_i(0, y, \gamma) = y f_{i2}(\gamma);$$

schließlich:

$$g_i(x, y, \gamma) = g_i(x, 0, \gamma) + g_i(0, y, \gamma) = x f_{i1}(\gamma) + y f_{i2}(\gamma); \quad (i = 1, 2).$$

Die Gleichungen

$$x'' = x f_{11}(\gamma) + y f_{12}(\gamma),$$

$$y'' = x f_{21}(\gamma) + y f_{22}(\gamma)$$

drücken die Abbildung der Ebene G_2 in sich aus, die durch das Element \mathcal{U}_γ der adjungierten Gruppe erzeugt wird. Wenn die Matrix (f_{ik}) nicht konstant ist, so bilden die den verschiedenen Werten von γ entsprechenden Abbildungen der Ebene G_2 in sich eine eingliedrige kontinuierliche Untergruppe der linearen homogenen Gruppe der Ebene.

Seien γ_1 und γ_2 zwei beliebige reelle Zahlen; die dem Wert $\gamma_1 + \gamma_2$ des Parameters entsprechende Matrix $((f_{ik}))$ ist das Produkt der Matrizen, die den Werten γ_1 und γ_2 des Parameters entsprechen:

$$((f_{ik}))_{\gamma_1 + \gamma_2} = ((f_{ik}))_{\gamma_1} \cdot ((f_{ik}))_{\gamma_2};$$

es folgt daraus, daß die Determinante:

$$\Delta(\gamma) = f_{11}(\gamma)f_{22}(\gamma) - f_{12}(\gamma)f_{21}(\gamma)$$

die Funktionalgleichung

$$\Delta(\gamma_1 + \gamma_2) = \Delta(\gamma_1) \cdot \Delta(\gamma_2)$$

befriedigt und als deren stetige Lösung sich durch die Formel

$$\Delta(\gamma) = e^{2a\gamma} \quad (a = \text{konst})$$

ausdrückt.

Führen wir die Funktionen φ_{ik} mittels der folgenden Gleichungen ein:

$$\varphi_{ik}(\gamma) = e^{-a\gamma} f_{ik}(\gamma) \quad (i, k = 1, 2).$$

Wenn die Matrix $((\varphi_{ik}))$ nicht konstant ist, setzen wir $t = x/y$; den Matrizen $((\varphi_{ik}))$ entsprechen durch eine stetige und einstufige Isomorphie die Transformationen

$$t' = \frac{\varphi_{11}(\gamma)t + \varphi_{12}(\gamma)}{\varphi_{21}(\gamma)t + \varphi_{22}(\gamma)},$$

welche eine eingliedrige kontinuierliche Untergruppe der Gruppe der linearen Transformationen von der Geraden t auf sich bilden. Eine solche Untergruppe ist mit einer der folgenden drei Gruppen identisch:

- a) Die Gruppe der parabolischen Transformationen mit einem gegebenen Fixpunkt t_0 ;
- b) die Gruppe der hyperbolischen Transformationen mit zwei gegebenen Fixpunkten t_1 und t_2 ;
- c) die Gruppe der elliptischen Transformationen mit zwei gegebenen komplex konjugierten Fixpunkten t_0 und \bar{t}_0 .

Die adjungierte Gruppe von G_3 läßt im Falle a) die Gerade $x/y = t_0$ und im Falle b) die Geraden $x/y = t_1, t_2$ invariant; im Falle c) gibt es keine durch den Anfangspunkt gehende Gerade in der Ebene, die bei der adjungierten Gruppe invariant ist.

Da für (S^*) und (T^*) zwei willkürliche eingliedrige Untergruppen von G_2 gewählt werden können, setzen wir voraus, daß in den Fällen a) und b) (S^*) eine invariante Gerade, und im Falle b) (T^*) die andere invariante Gerade der adjungierten Gruppe ist. Im Falle c) wählen wir die Untergruppen (S^*) und (T^*) derart, daß die Minimalgeraden $z \pm iy = 0$ die beiden invarianten Geraden der adjungierten Gruppe in der Ebene G_2 darstellen. Durch diese Wahl der

Koordinaten (x, y) vereinfacht sich die Form der linearen Transformationen in t , und dementsprechend der Ausdruck der Matrix $((\varphi_{ik}))$ folgenderweise:

$$a) t' = t + \gamma; \quad \varphi_{11}(\gamma) = \varphi_{22}(\gamma) = 1, \quad \varphi_{12}(\gamma) = \gamma, \quad \varphi_{21}(\gamma) = 0;$$

$$b) t' = \frac{e^{\gamma} t}{e^{-\gamma}}; \quad \varphi_{11}(\gamma) = e^{\gamma}, \quad \varphi_{12}(\gamma) = \varphi_{21}(\gamma) = 0, \quad \varphi_{22}(\gamma) = e^{-\gamma};$$

$$c) t' = \frac{\cos \gamma \cdot t - \sin \gamma}{\sin \gamma \cdot t + \cos \gamma}; \quad \varphi_{11}(\gamma) = \varphi_{22}(\gamma) = \cos \gamma, \quad -\varphi_{12}(\gamma) = \varphi_{21}(\gamma) = \sin \gamma.$$

Außer den obigen Matrizen müssen wir auch die konstante Matrix:

$$\varphi_{11}(\gamma) = \varphi_{22}(\gamma) = 1, \quad \varphi_{12}(\gamma) = \varphi_{21}(\gamma) = 0$$

berücksichtigen.

Durch Multiplizieren mit $e^{a\gamma}$ erhalten wir aus den Matrizen $((\varphi_{ik}))$ die $((f_{ik}))$. Jeder Matrix $((f_{ik}))$, die auf diese Weise erhalten wird, entspricht je eine eingliedrige Untergruppe der linearen homogenen Gruppe der Ebene. Wir erhalten somit das folgende Ergebnis:

Satz 3. Jede eingliedrige kontinuierliche Untergruppe der linearen homogenen Gruppe der Ebene läßt sich durch geeignete Wahl der Koordinaten mittels einer der folgenden Formeln ausdrücken:

$$\begin{array}{ll} 1. x' = e^{\gamma} x, & y' = e^{\gamma} y; \\ 2. x' = x + \gamma y, & y' = y; \\ 3. x' = e^{a\gamma} x + e^{a\gamma} \gamma y, & y' = e^{a\gamma} y; \\ 4. x' = e^{(a+1)\gamma} x, & y' = e^{(a-1)\gamma} y; \\ 5. x' = x \cos \gamma - y \sin \gamma, & y' = x \sin \gamma + y \cos \gamma; \\ 6. x' = e^{a\gamma} (x \cos \gamma - y \sin \gamma), & y' = e^{a\gamma} (x \sin \gamma + y \cos \gamma). \end{array}$$

Wir bemerken, daß sich das System der Funktionalgleichungen:

$$((f_{ik}))_{\gamma_1 + \gamma_2} = ((f_{ik}))_{\gamma_1} \cdot ((f_{ik}))_{\gamma_2},$$

welches zur Bestimmung der eingliedrigen Untergruppen der linearen homogenen Gruppe dient, durch elementare und sehr einfache Rechnungen auflösen läßt, ähnlicherweise, wie das System der Funktionalgleichungen zur Bestimmung der Funktionen $\sin x$ und $\cos x$ ⁵⁾. Wir haben auf ein solches rechnerisches Verfahren verzichtet und eine Methode angewendet, die auf dem Isomorphismus mit den Untergruppen der linearen Gruppe der Geraden beruht. Der Grundgedanke dieser Methode rührt von S. Lie her⁶⁾.

Aus dem Satz 3 folgt laut der zu Anfang dieses Paragraphen gegebenen Ausführungen:

⁵⁾ Siehe z. B. W. F. Osgood, Lehrbuch der Funktionentheorie, Bd. I (2. Aufl., Leipzig und Berlin, 1912), S. 582.

⁶⁾ S. Lie und G. Scheffers, Vorlesungen über kontinuierliche Gruppen (Leipzig, 1893), S. 134–149.

Satz 4. Die dreigliedrige integrierbare Gruppe G_3 ist einer der folgenden Gruppen homöomorph:

- | | | |
|-----------------------------------------------------------------|-------------------------------------------------------------|--------------------|
| 1. $x' = e^x x + \alpha,$ | $y' = e^x y + \beta,$ | $z' = z + \gamma;$ |
| 2. $x' = x + \gamma y + \alpha,$ | $y' = y + \beta,$ | $z' = z + \gamma;$ |
| 3. $x' = e^{a\gamma} x + e^{a\gamma} \gamma y + \alpha,$ | $y' = e^{a\gamma} y + \beta,$ | $z' = z + \gamma;$ |
| 4. $x' = e^{(a+1)\gamma} x + \alpha,$ | $y' = e^{(a-1)\gamma} y + \beta,$ | $z' = z + \gamma;$ |
| 5. $x' = x \cos \gamma - y \sin \gamma + \alpha,$ | $y' = x \sin \gamma + y \cos \gamma + \beta,$ | $z' = z + \gamma;$ |
| 6. $x' = e^{a\gamma} (x \cos \gamma - y \sin \gamma) + \alpha,$ | $y' = e^{a\gamma} (x \sin \gamma + y \cos \gamma) + \beta,$ | $z' = z + \gamma;$ |
| 7. $x' = x + \alpha,$ | $y' = y + \beta,$ | $z' = z + \gamma.$ |

Als Korollar von diesem Satz ergibt sich der folgende:

Satz 5. Jede nicht-kommutative dreigliedrige integrierbare Gruppe läßt sich als eine Gruppe von Affinitäten der Ebene verwirklichen.

Unter den Gruppen, die im Satz 4 von 1 bis 6 aufgezählt worden sind, besitzt nur die Gruppe 5, nämlich die Gruppe der Bewegungen der euklidischen Ebene, die Eigenschaft, daß sämtliche in ihr enthaltenen Abbildungen in der (x, y) -Ebene überall regulär⁷⁾ sind. Aus den Sätzen 4 und 5 ergibt sich also die folgende neue Charakterisierung der euklidischen ebenen Bewegungsgruppe:

Satz 6. Jede dreigliedrige integrierbare Gruppe von regulären Abbildungen der Ebene in sich ist der Bewegungsgruppe der euklidischen Ebene homöomorph.

Aus dem Satz 4 ergibt sich ferner die folgende rein gruppentheoretische Charakterisierung der euklidischen ebenen Bewegungsgruppe:

Satz 7. Wenn G_3 eine nicht-kommutative dreigliedrige integrierbare Gruppe ist, und wenn die Faktorgruppe G_3/G_2 von G_3 in bezug auf eine zweigliedrige invariante Untergruppe G_2 zyklisch ist, so ist G_3 der Parametergruppe der euklidischen ebenen Bewegungsgruppe homöomorph.

Bemerken wir zum Schluß, daß, wenn die in G_3 enthaltene zweigliedrige invariante Untergruppe G_2 der Gruppe der Bewegungen eines Zylinders in sich homöomorph ist, der Ausdruck von G_3 aus den Formeln 2 oder 7 des Satzes 4 durch Reduktion von x mod 1, und wenn G_2 der Gruppe der Verschiebungen des Torus homöomorph ist, aus der Formel 7 des Satzes 4 durch Reduktion von x und von y mod 1 erhalten wird.

Die vorliegende Arbeit wurde der Ungarischen Akademie der Wissenschaften in der Sitzung vom 19. Mai 1941 vorgelegt.

⁷⁾ Vgl. B. von Kerékjártó, Topologische Charakterisierung der linearen Abbildungen. Acta Scient. Math. Szeged 6 (1934), S. 235–262.

Über verstreute Mengen.

Von

Georg v. Alexits in Budapest.

I. Man kann die verstreuten Punktmengen in vier Klassen einteilen: 1. Die Klasse \mathfrak{R}_1 der *diskontinuierlichen* Mengen, zu welcher die Mengen ohne Teilkontinua gehören. 2. Die Klasse \mathfrak{R}_2 der *zusammenhanglosen* Mengen, zu welcher die Mengen mit einpunktigen Komponenten gehören. 3. Die Klasse \mathfrak{R}_3 der *total zusammenhanglosen* Mengen, zu welcher die Mengen mit einpunktigen Quasikomponenten gehören, was so viel bedeutet, daß eine Menge aus \mathfrak{R}_3 für je zwei ihrer Punkte p, q in zwei getrennte Mengen zerlegt werden kann, deren eine den Punkt p , die andere den Punkt q enthält. 4. Die Klasse \mathfrak{R}_4 der nulldimensionalen Mengen. Diese sind Mengen, deren Punkte in beliebig kleinen Umgebungen mit zur Menge fremden Begrenzungen enthalten sind.

Diese vier Mengenklassen sind im allgemeinen verschieden, und zwar ist $\mathfrak{R}_4 \subset \mathfrak{R}_3 \subset \mathfrak{R}_2 \subset \mathfrak{R}_1$; unter Umständen können sie aber zusammenfallen, z. B. gilt $\mathfrak{R}_1 = \mathfrak{R}_2 = \mathfrak{R}_3 = \mathfrak{R}_4$ für die Teilmengen einer Geraden, oder auch für die Teilmengen der sogenannten Bäume¹⁾, d. i. Kontinua, die keinen topologischen Kreis enthalten. Es wäre von gewissem dimensionstheoretischen Interesse, die Kontinua zu kennzeichnen, in welchen diese vier Klassen zusammenfallen²⁾. Die diesbezüglichen Untersuchungen sind aber noch durchaus nicht abgeschlossen; die bisher erzielten allgemeinsten Ergebnisse scheinen die folgenden zu sein:

I. $\mathfrak{R}_1 = \mathfrak{R}_2 = \mathfrak{R}_3 = \mathfrak{R}_4$ gilt für die Teilmengen der Bäume¹⁾.

II. $\mathfrak{R}_2 = \mathfrak{R}_3 = \mathfrak{R}_4$ gilt für die Teilmengen der regulären Kurven³⁾.

III. $\mathfrak{R}_2 = \mathfrak{R}_3$ gilt für die Teilmengen der erblich lokal zusammenhängenden Kontinua⁴⁾.

¹⁾ G. T. Whyburn, Monatsh. Math. Phys. 38 (1931), S. 85; vgl. auch G. v. Alexits, ebenda 33 (1933), S. 407.

²⁾ Vgl. zu dieser Fragestellung K. Menger, Kurventheorie, S. 369 (Leipzig und Berlin 1932). Wir schließen uns übrigens in der Terminologie diesem Standardwerk der mengentheoretischen Kurventheorie an. Für die Ausdrücke: reguläre, rationale, irrationale Kurve und Ordnung eines Punktes vgl. insbesondere S. 96–98.

³⁾ G. T. Whyburn, A. a. O.¹⁾

⁴⁾ G. T. Whyburn, Amer. Journ. Math. 53 (1931), S. 374. Lokal zusammenhängend bedeutet dabei, daß jeder Punkt einer vorgegebenen Menge beliebig kleine Umgebungen hat, deren Durchschnitte mit der betrachteten Menge zusammenhängend sind. Erblich lokal zusammenhängend heißt ein Kontinuum, dessen sämtliche Teilkontinua lokal zusammenhängend sind.

Es soll nun gezeigt werden, daß II einer starken Erweiterung fähig ist, durch welche zugleich auch III verallgemeinert wird. Wir werden nämlich den folgenden Satz beweisen:

IV. $\mathfrak{R}_2 = \mathfrak{R}_3 = \mathfrak{R}_4$ gilt für die Teilmengen der Summen abzählbar vieler erblich lokal zusammenhängender Kontinua.

Auf Grund dieses Satzes kann man auch zwei Sätze⁵⁾ über reguläre Kurven auf die bedeutend allgemeinere Klasse der Kontinua, welche sich als Summen abzählbar vieler erblich lokal zusammenhängender Kontinua darstellen lassen, übertragen.

2. Satz IV ist bewiesen, wenn wir zeigen, daß eine nicht zu \mathfrak{R}_4 gehörende Teilmenge A einer abzählbaren Summe erblich lokal zusammenhängender Kontinua nicht zu \mathfrak{R}_2 gehören kann, was so viel bedeutet, daß die Beziehungen

$$(1) \quad \dim A > 0,$$

$$(2) \quad A \in \mathfrak{R}_2$$

miteinander unverträglich sind. Wir nehmen zum Beweise unserer Behauptung das Gegenteil an und leiten aus dieser Annahme einen Widerspruch her.

Es seien also $K_1, K_2, \dots, K_n, \dots$ abzählbar viele erblich lokal zusammenhängende Kontinua. Ferner bezeichne A eine Teilmenge der Summe $\sum_{n=1}^{\infty} K_n$, welche sowohl der Beziehung (1) wie auch der Beziehung (2) genügt.

Wegen $A = \sum_{n=1}^{\infty} AK_n$ kann der Durchschnitt AK_n nicht für alle $n = 1, 2, \dots$ nulldimensional sein, da sonst auch A als Summe der abzählbar vielen in A abgeschlossenen Mengen AK_n nulldimensional sein sollte⁶⁾, gegen die angenommene Beziehung (1). Es gibt also einen Index n_0 , für welchen AK_{n_0} nicht nulldimensional ist. Da die Menge A nach (2) zusammenhangslos ist, ist um so mehr ihr Teil AK_{n_0} zusammenhangslos. Daraus folgt nach II, daß die Menge AK_{n_0} sogar total zusammenhangslos ist. Beachten wir, daß nach einem Mengerschen Satz⁷⁾ die nulldimensionalen Mengen unter den total zusammenhangslosen Teilen eines kompakten metrischen Raumes dadurch gekennzeichnet sind, daß sie auch nach Hinzufügung eines beliebigen Punktes total zusammenhangslos bleiben, so folgt, da K_{n_0} als Raum betrachtet kompakt ist und seine Teilmenge AK_{n_0} , — wie es soeben gezeigt wurde — nicht null-

⁵⁾ G. de Alexits, C. R. Soc. Sci. Lett. Varsovie 31 (1939), S. 104.

⁶⁾ W. Hurewicz, Math. Annalen 96 (1927), S. 760.

⁷⁾ K. Menger, Ber. Wiener Akad. 133 (1924), S. 421.

dimensional ist, die Existenz eines Punktes $p \in K_{n_0}$ derart, daß die Menge $AK_{n_0} + (p)$ nicht mehr total zusammenhangslos ist. Nach II enthält dann $AK_{n_0} + (p)$ eine mehrpunktige Komponente B . Bezeichne q einen von p verschiedenen Punkt der Menge B und sei U_q eine p nicht enthaltende Umgebung von q ⁸⁾. Weil B ein mehrpunktiger zusammenhängender Teil des erblich lokal zusammenhängenden Kontinuums K_{n_0} ist, folgt aus einem Wilderschen Satz⁹⁾ über erblich lokal zusammenhängende Kontinua, daß die Menge B lokal zusammenhängend ist, mithin enthält U_q eine Umgebung V_q von q , deren Durchschnitt mit B zusammenhängend ist. Es besteht aber die Beziehung

$$BV_q \subset BU_q \subset B - (p) \subset AK_{n_0},$$

wo BV_q zusammenhängend und als offener Teil der mehrpunktigen zusammenhängenden Menge B mehrpunktig ist. Die Menge AK_{n_0} enthält also einen mehrpunktigen zusammenhängenden Teil BV_q , sie ist m. a. W. nicht zusammenhangslos, entgegen der angenommenen Beziehung (2). Damit ist der angekündigte Widerspruch hergestellt und Satz IV bewiesen.

3. Aus IV läßt sich folgende Verallgemeinerung eines vom Verfasser für reguläre Kurven bewiesenen Satzes⁵⁾ herleiten:

V. Ist A ein positiv dimensionales absolutes Teil- G_δ oder F_σ einer abzählbaren Summe erblich lokal zusammenhängender Kontinua, so enthält A einen Bogen.

Sei zunächst A ein Teil- G_δ der Summe $\sum_{n=1}^{\infty} K_n$, wo die K_n erblich lokal zusammenhängende Kontinua bezeichnen. Wie beim Beweise des Satzes IV folgt aus $\dim A > 0$ die Existenz eines n_0 , so daß $\dim AK_{n_0} > 0$. Nach IV enthält AK_{n_0} eine mehrpunktige zusammenhängende Menge B . Betrachten wir nun den Durchschnitt $AK_{n_0} \cdot \bar{B}$ etwas näher, wo mit \bar{B} die abgeschlossene Hülle der Menge B bezeichnet wurde: Weil A , K_{n_0} und \bar{B} absolute G_δ -Mengen sind, ist auch ihr Durchschnitt $AK_{n_0} \cdot \bar{B}$ ein absolutes G_δ , und zwar ein zusammenhängendes absolutes G_δ , da $B \subset AK_{n_0}$ zusammenhängend ist. Nach dem soeben erwähnten Satze des Herrn Wilder⁹⁾ folgt daraus, daß der Durchschnitt $AK_{n_0} \cdot \bar{B}$ lokal zusammenhängend ist. Die Menge $AK_{n_0} \cdot \bar{B}$ hat somit die Eigenschaft, ein mehrpunktiges, zusammenhängendes, lokal zusammen-

⁸⁾ Wir betrachten K_{n_0} als metrischen Raum, die Umgebung U_q ist also als „Umgebung im Raume K_{n_0} “ aufzufassen.

⁹⁾ R. L. Wilder, Proc. Nat. Ac. Sci. U. S. A. 15 (1929). S. 616.

hängendes absolutes G_j zu sein, woraus sich nach einem Mengerschen Satz¹⁰⁾ die Existenz eines in $A K_{n_0} \cdot \bar{B}$ enthaltenen Bogens C ergibt. Damit ist unsere Behauptung bezüglich der G_j bewiesen.

Sei sodann A ein absolutes F_n . Wegen $\dim A > 0$ enthält A nach einem Satze des Herrn Mazurkiewicz¹¹⁾ ein Kontinuum. Dieses Kontinuum ist ein positiv dimensionales absolutes Teil- G_j der Menge A , folglich enthält es laut des soeben bewiesenen ersten Teiles unserer Behauptung einen Bogen, womit auch der zweite Teil bewiesen ist.

4. Wir nennen eine Teilmenge M eines Raumes R *höchstens beständig nulldimensional*, wenn die Summe von M und einer beliebigen nulldimensionalen Teilmenge von R nulldimensional ist. Als Verallgemeinerung eines vom Verfasser zunächst für Bäume¹⁾, später für reguläre Kurven⁵⁾ bewiesenen Satzes läßt sich der folgende Satz beweisen:

VI. *Sei K ein Teilkontinuum einer Summe abzählbar vieler erblich lokal zusammenhängender Kontinua, so ist die Menge K^1 der Punkte erster Ordnung (Endpunkte) von K höchstens beständig nulldimensional.*

Denn im entgegengesetzten Fall läßt sich eine nulldimensionale Teilmenge M von K bestimmen, so daß $\dim(K^1 + M) > 0$. In K gibt es aber bekanntlich¹²⁾ auch eine M enthaltende nulldimensionale absolute G_j -Menge N ; dann gilt

$$\dim(K^1 + N) \geq \dim(K^1 + M) > 0.$$

Da sowohl N als auch K^1 absolute Teil- G_j von K sind¹³⁾, ist die Menge $K^1 + N$ ein positiv dimensionales absolutes Teil- G_j der K enthaltenden Summe von abzählbar vielen erblich lokal zusammenhängenden Kontinua. $K^1 + N$ enthält also nach V einen Bogen C . Der Durchschnitt CK^1 besteht offensichtlich aus lauter Punkten erster Ordnung des Bogens C , d. h. aus den beiden Endpunkten von C . Dann enthält aber $CN = C - CK^1$ einen Teilbogen von C , folglich ist CN nicht nulldimensional, gegen die Annahme $\dim N = \dim M = 0$. Aus diesem Widerspruch folgt die Unmöglichkeit der angenommenen Beziehung $\dim(K^1 + M) > 0$, d. h. die Menge K^1 ist höchstens beständig nulldimensional, wie behauptet.

¹⁰⁾ K. Menger, Monatsh. Math. Phys. 36 (1929), S. 212; C. Kuratowski, Fund. Math. 15 (1930), S. 301; vgl. auch N. Aronszajn, Fund. Math. 15 (1930), S. 228.

¹¹⁾ S. Mazurkiewicz, Fund. Math. 3 (1922), S. 65.

¹²⁾ A. a. O.⁴⁾ und L. Tumarkin, Math. Annalen 98 (1928), S. 638.

¹³⁾ A. a. O.³⁾, S. 104.

5. Satz VI kann nicht dahin verschärft werden, daß die Menge K^1 der Punkte erster Ordnung einer beliebigen Kurve höchstens beständig nulldimensional ist. In der Tat, es existiert eine schwach irrationale Kurve K , welche Summe abzählbar unendlich vieler Strecken $S_1, S_2, \dots, S_n, \dots$ und der Menge K^1 der Punkte erster Ordnung von K ist¹⁴⁾. Die Menge $\sum_{n=1}^{\infty} S_n$ läßt sich bekanntlich als Summe $A + B$ zweier Mengen darstellen, von welchen A abzählbar und B nulldimensional ist¹⁵⁾. Die Kurve K läßt sich also folgenderweise darstellen:

$$K = A + B + K^1.$$

Wäre K^1 höchstens beständig nulldimensional, so wäre die Menge $B + K^1$ nulldimensional. Das bedeutet, daß jeder Punkt von $B + K^1$ beliebig kleine Umgebungen hätte, deren Begrenzungen nicht zu $B + K^1$, sondern zur abzählbaren Menge A gehören, d. h. jeder Punkt von $B + K^1$ wäre rational. Die Kurve K könnte mithin höchstens in den Punkten der nulldimensionalen Menge A irrational sein, was aber unmöglich ist, da die Menge der irrationalen Punkte nicht nulldimensional sein kann¹⁶⁾. Die Annahme $\dim(B + K^1) = 0$ führt also zu einem Widerspruch, d. h. die Menge K^1 ist nicht höchstens beständig nulldimensional.

Beachtet man, daß unser Beweis der nicht beständigen Nulldimensionalität von K auf die Irrationalität der Knaster-Mazurkiewicz'schen Kurve aufgebaut war, so ist man geneigt, folgende Frage zu stellen: Ist K^1 für jede rationale Kurve höchstens beständig nulldimensional?

6. Ein Kontinuum K heiße erblich lokal zusammenhängend im Punkte p , wenn p eine Umgebung hat, deren abgeschlossene Hülle ein erblich lokal zusammenhängendes Teilkontinuum von K ist. Zum Schluß beweisen wir noch folgende Struktureigenschaft der Summen erblich lokal zusammenhängender Kontinua:

VII. Ist das Kontinuum K Summe der abzählbar vielen erblich lokal zusammenhängenden Kontinua $K_1, K_2, \dots, K_n, \dots$, so ist die Menge der Punkte, in welchen K nicht erblich lokal zusammenhängend ist, nirgends dicht in K .

Sei in der Tat U eine beliebige offene Menge des als Raum betrachteten Kontinuums $K = \sum_{n=1}^{\infty} K_n$. Wäre jede Menge $K_n U$ nirgends dicht, so würde

$U = \sum_{n=1}^{\infty} K_n U$ von erster Kategorie sein, was aber nicht der Fall ist, da U als

¹⁴⁾ B. Knaster und S. Mazurkiewicz, *Ergebn. Math. Koll. Wien* 5 (1933), S. 7.

¹⁵⁾ K. Menger, *Dimensionstheorie*, S. 120 (Leipzig und Berlin 1928).

¹⁶⁾ A. a. O.²⁾, S. 183.

offener Teil des Kontinuums K von zweiter Kategorie sein muß. Es gibt also einen Index n_0 , so daß $K_{n_0}U$ nicht nirgends dicht ist. Daraus folgt, weil $K_{n_0}U$ abgeschlossen in U ist, daß $K_{n_0}U$ eine offene Teilmenge V von K enthält. Sei p ein Punkt von V . Wegen $V \subset K_{n_0}$ und des erblich lokalen Zusammenhanges des Kontinuums K_{n_0} gibt es eine Umgebung U_p von p derart, daß ihre abgeschlossene Hülle \bar{U}_p ein erblich lokal zusammenhängendes Teilkontinuum von V ist. Jeder Punkt q der offenen Menge V gehört mithin zur Menge M der Punkte, in welchen K erblich lokal zusammenhängend ist. Die beliebig gewählte offene Teilmenge U von K enthält also einen nicht leeren offenen Teil der Menge M , d. h. der offene Kern M_i von M ist dicht in K . Dann ist aber seine abgeschlossene Komplementärmenge $K - M_i$ nirgends dicht in K , um so mehr ist also wegen $K - M \subset K - M_i$ die Menge $K - M$ der Punkte, in welchen K nicht erblich lokal zusammenhängend ist, nirgends dicht in K .

(Eingegangen am 12. 2. 1941.)

Parametrixmethode zur Lösung von Randwertproblemen*).

Von

Ludwig Sauer in Frankfurt am Main.

Teil 1.

Auflösung von Integralgleichungen zweiter Art.

Vorbemerkung.

a. In der Literatur löst man vorgelegte Randwertprobleme meistens mit Hilfe der *Greenschen Funktion* G . Diese ist für das ebene Δ -Problem und eine Kreisfläche als Grundbereich bekannt, wenn die gesuchte Funktion auf dem Kreisumfang verschwinden soll. Ferner wird G in der Literatur auch für weitere einfache Probleme unmittelbar angegeben. G für allgemeinere Differentialausdrücke und Grundbereiche führt man schrittweise auf die schon bekannten G zurück. Dabei setzt man voraus:

1. Die Angabe der Transformation eines vorgelegten allgemeinen Differentialausdrucks auf die Normalform.

2. Das Auffinden der konformen Abbildung des Grundbereichs auf einen Bereich, für den G im Δ -Fall bekannt ist.

3. Die Auflösung einer Integralgleichung, welche beispielsweise dem Randwertproblem in der Normalform gleichwertig ist, und auf die dieses Problem mittels des für das Δ -Problem aufgestellten G zurückgeführt wird. Danach findet man mittelbar das G für das allgemeinere Problem.

4. Die Konvergenz einer unendlichen Folge von G für gewisse einfachere Probleme mit eckenfreien Grundbereichen nach dem G für das Ausgangsproblem mit eckigem Grundbereich; dabei nähern sich die eckenfreien Bereiche in bestimmter Weise dem Ausgangsbereich.

Meistens werden diese Schritte mehrfach zusammengesetzt und ergeben so ein kompliziertes und indirektes Verfahren zur Bildung von G für das Ausgangsproblem vor allem dann, wenn der Grundbereich Ecken hat.

Verzichtet man jedoch vorerst auf die Darstellung der Lösung des Problems als Integral über das Produkt von G und der rechten Seite der Differentialgleichung, so genügt es zur Auflösung des Problems, G für den reduzierten Differentialausdruck, etwa für die Summe der Glieder mit den Ableitungen höchster Ordnung, zu finden. Dann läßt sich das Ausgangsproblem mit diesem reduzierten G auf eine ihm *gleichwertige* lösbare Integralgleichung zweiter

*) D 30.

Art mit uneigentlich singulärem, unsymmetrischem Kern zurückführen (vgl. obiges 3). Die allgemeine Lösung der Integralgleichung löst auch das Randwertproblem. Ferner läßt sich aus dem reduzierten G das vollständige G und damit schließlich die schon erwähnte Lösungsdarstellung angeben.

G ist schwierig zu bilden, weil G vor allem folgende Eigenschaften haben muß, wobei ich unter t den veränderlichen Punkt und unter t' den Parameterpunkt verstehe:

a) $G = G(t, t')$ soll der vorgegebenen oder einer reduzierten Differentialgleichung im Innern des Grundbereiches bei festem t' als Funktion von t für $t \neq t'$ genügen.

b) G soll die vorgegebene Randbedingung für $t \neq t'$ erfüllen.

c) G soll für $t = t'$ eine gewisse Singularität besitzen.

b. Der Umweg der Auflösung des Randwertproblems über das angedeutete Bildungsverfahren für G läßt sich durch eine grundsätzlich einfachere, unmittelbare Auflösungsmethode vermeiden. Um nämlich das Randwertproblem auf eine Integralgleichung zurückzuführen, ist die Eigenschaft a) unnötig. Es genügt an deren Stelle die allgemeinere, daß die mittels den Eigenschaften b) und c) abgeleitete Integralgleichung lösbar ist. Eine Funktion, die das leistet, heiße *Parametrix* $P = P(t, t')$. Sie führt auf eine dem Problem *ungleichwertige* Integralgleichung, aus deren allgemeinen Lösung erst diejenige des Randwertproblems ausgesondert werden muß. Diesen neben b) und c) tretenden Forderungen an P , nämlich die Lösbarkeit der Integralgleichung und die Aussonderung der Lösung des Randwertproblems aus derjenigen der Integralgleichung zu ermöglichen, werde ich unter Vermeidung obiger für G angegebenen Schritte 1 bis 4 im Abschnitt II durch eine unmittelbare Bildung von P genügen. Letztere ist elementar gegenüber der im allgemeinen komplizierten Bildung von G , was gerade in dem Wegfall der Eigenschaft a) für P begründet ist. Auf die genauen Eigenschaften von P , insbesondere auf die Stetigkeits- und Differenzierbarkeitseigenschaften, sowie auf die Symmetrie $P(t, t') = P(t', t)$ gehe ich im Abschnitt III im einzelnen ein. Da P doch nicht einer Differentialgleichung genügt, ist das für ein reduziertes Randwertproblem anzugebende P damit auch eine Parametrix für das ursprüngliche Problem (und umgekehrt). Mittels P wird die Auflösungstheorie für das vorgelegte Problem entwickelt. Weitergehend läßt sich aus P eine Lösungsdarstellung und damit schließlich G bilden.

c. Hilbert wies auf eine solche Möglichkeit hin und entwickelte diese „neue Methode an dem Beispiel der allgemeinsten linearen partiellen Differentialgleichung vom elliptischen Typus, während das Integrationsgebiet die Vollkugel ist“. Letztere ist die geschlossene Kugeloberfläche, wodurch die eigentliche Randbedingung wegfällt.

Diese P -Methode für elliptische Differentialgleichungen läßt sich mittels der im Abschnitt II zu bildenden Parametrix auf Randwertprobleme ausdehnen und zwar auf allgemeinere Klassen von eigentlichen linearen Randwertproblemen erster Art der reellen Analysis. Die diesbezüglichen Ausführungen bleiben späteren Veröffentlichungen vorbehalten. Hierdurch werden nicht nur — wie im Abschnitt IV — die Lösungsverfahren bekannter Probleme durch einfachere ersetzt, sondern auch — wie im Abschnitt V — solche für neue Probleme hinzugewonnen. In Sonderfällen wird P so elementar gebildet, daß damit vielleicht eine numerische Behandlung eingeleitet werden könnte. Stets wurde versucht, die Herausarbeitung der P -Methode und ihrer Begründung mit der genauen Angabe ihrer Ergebnisse und deren Voraussetzungen in Einklang zu bringen. Die Lösungsformeln treten gegenüber den Lösungstatsachen zurück. Das sind die bekannten Hauptsätze der Auflösungstheorie (Abschnitt IV), die in weitergehenden Sätzen über das allgemeinste G eine Ergänzung erhalten. Letzteres bekommt auch dadurch eine Bedeutung, daß G aus P gebildet wird.

d. Die Auflösungssätze werden zu denjenigen in Parallele gesetzt, die für die genannte *Integralgleichung* gelten, auf die das Randwertproblem mittels P zurückgeführt wird. Die hierzu benötigte Integralgleichungstheorie wird über das zur Anwendung bei der P -Methode Erforderliche hinausgehend verallgemeinert und als ein selbständiger Beitrag zur Literatur der P -Methode als Abschnitt I vorangestellt. Hierbei handelt es sich nicht um Berichte über bekannte Sätze, sondern erstens um die Vervollständigung der Beweise der Hauptsätze der Auflösungstheorie für zwei wichtige Klassen von uneigentlich singulären Integralgleichungen und dazu Verallgemeinerung bekannter Hilfsätze. Zweitens werden weitere Eigenschaften des allgemeinsten lösenden Kernes festgestellt.

Literatur.

Courant-Hilbert.

Methoden der math. Physik, Bd. 1 (1924); Bd. 2 (1937).

Frank-Mises.

Die Differential- und Integralgleichungen der Mechanik und Physik (1925—1927).

Fredholm.

Sur une classe d'équations fonctionnelles. Acta mathematica 27, S. 365—390, Stockholm 1903.

Hahn-Lichtenstein-Lense.

Integralgleichungen und ihre Anwendungen auf Randwertaufgaben. Artikel in Pascals Repertorium der höheren Analysis, Bd. I, 3, 2. Aufl., (1929).

Haupt.

Zur Parametrixmethode. Math. Annalen 88 (1922), Heft 1—2, S. 136—150.

Hellinger.

Hilberts Arbeiten über Integralgleichungen und unendliche Gleichungssysteme. Art. IX von Hilbert, Gesammelte Abh., Bd. III (1935).

Hellinger-Toeplitz.

Integralgleichungen und Gleichungen mit unendlichvielen Unbekannten. Enzykl.-Art. II, C 13 (1927).

Hilbert.

Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, vor allem Kap. XVIII: Eine neue Methode der Zurückführung von Differentialgleichungen auf Integralgleichungen. — Begriff der Parametrix. Gött. Nachr. (1910), 6. Mitt., S. 8–34.

Iglsch.

Die determinantepfreien Sätze bei linearen Integralgleichungen. Math. Annalen 110 (1934), 2. Heft, S. 223–229.

E. E. Levi.

Sulle equazioni integrali. Atti della reale accademia dei Lincei, serie 5, Rendiconti, Volume XVI, 2. Rom 1907. S. 604–612.

Lichtenstein.

Neuere Entwicklung der Theorie partieller *Differentialgleichungen* 2. Ordnung vom elliptischen Typus. Enzykl.-Art. II, C 12 (1924).

Lichtenstein über *konforme Abbildung*:

Beweis des Satzes, daß jedes hinreichend kleine, im wesentlichen stetig gekrümmte, singularitätenfreie Flächenstück auf einen Teil einer Ebene zusammenhängend und in den kleinsten Teilen ähnlich abgebildet werden kann. Verlag der kgl. Akad. der Wiss., Berlin 1912.

Erhardt Schmidt.

Zur Theorie der linearen und nichtlinearen Integralgleichungen. Math. Annalen 1907.

Teil I: Entwicklung willkürlicher Funktionen nach Systemen vorgeschriebener. Bd. 63, S. 433–476.

Teil II: Auflösung der allgemeinen linearen Integralgleichungen. Bd. 64, S. 161–174.

Sommerfeld.

Randwertaufgaben in der Theorie der partiellen Differentialgleichungen. Enzykl.-Art. II, A 7 c (1900).

Sternberg.

Die Theorie der Randwertaufgaben im Gebiete der partiellen Differentialgleichungen. Art. in Pascals Repertorium der höheren Analysis, Bd. I 3, 2. Aufl. (1929).

Am Schluß der Veröffentlichungsreihe über die Parametrixmethode wird ein *Inhaltsverzeichnis* mit Seitenangaben gebracht.

I. Auflösungstheorie der Integralgleichungen zweiter Art auf Grund von Kerniteration und Kernabspaltung.

1. Fredholmsche und Levische Integralgleichungen.

Zuerst gebe ich die allgemein zu lösenden Integralgleichungen an.

a. Die unabhängigen Veränderlichen seien x_v für $v = 1 \dots q$, die ich im Fall $q = 2$ meist mit x und y bezeichne; q sei eine endliche, ganze Zahl größer als eins. Die x_v deute ich als rechtwinklige Koordinaten eines Punktes t im q -dimensionalen Zahlenraum. Der Ausgangsbereich b für t sei reell, q -dimensional, zusammenhängend, schlicht und abgeschlossen (beschränkt). (Vgl. Literatur „Hellinger-Toeplitz ...“, S. 1389.) t' sei ein zweiter Punkt. $K = K(t, t')$ und $M = M(t)$ seien vorgegebene Funktionen. $W = W(t)$ sei eine gesuchte Funktion. Man setze $K'' = K(t', t)$, $M' = M(t')$ und $W' = W(t')$.

Ich gehe in diesem Abschnitt I von der *allgemeinen Integralgleichung zweiter Art* aus, nämlich von

$$(1.1) \quad W - \int K \cdot W' dt' = M \text{ in } b,$$

wobei ich dt' über b integriere. Hierin wird angenommen:

I. Es existiere jedes Glied der Reihe

$$\int K \cdot M' dt, \quad \iint K(t, s) \cdot K(s, t') \cdot M' dt' ds; \\ \iiint K(t, r) \cdot K(r, s) \cdot K(s, t') \cdot M' dt' ds dr \text{ usw.}$$

Ferner existiere jedes Glied der entsprechenden Reihen, die mit $\int K \cdot W' dt'$, mit $\int K'' \cdot M' dt'$ und mit $\int K'' \cdot W' dt'$ beginnen. Schließlich existiere das Integral über dem Produkt von M oder W und irgendeinem der genannten Reihenglieder, ferner von M oder W und einer Funktion der gleichen Art wie M und W . Jedoch fordere ich nicht $K = K''$.

II. In jedem in I genannten mehrfachen Integral soll die Integrationsfolge vertauschbar sein. Dann sind die $(v-1)$ -ten iterierten Kerne $K^{(v)} = K^{(v)}(t, t')$ für $v = 2, 3, \dots$ vorhanden (Hell.-Toepl., S. 1383). $K^{(1)}$ bedeutet K . Es ist $K^{(v)(v)} = K^{(v)''}$.

III. Zu jedem in I auftretenden Integral, das eine Funktion von t ist, möge ein Teilbereich B von b der gleichen Art wie b mit dem Flächeninhalt f bestimmbar sein, der alle Unstetigkeitsstellen des Integranden enthält. B kann auch aus getrennten Stücken bestehen. Dann soll es möglich sein, f beliebig klein zu machen, ferner einer willkürlich vorgegebenen, positiven Zahl z eine Zahl $g(z)$ stets so zuzuordnen, daß der absolute Wert des über B genommenen Teilintegrals kleiner als z für $f < g(z)$ wird, und zwar gleichmäßig für t in b .

Anmerkung. Ich suche demnach nur solche Lösungen W , die durch vorstehende allgemeinen Eigenschaften eingeschränkt werden. Daß es solche W gibt, wird in diesem Abschnitt bewiesen.

b. Die Integralgleichung (1.1) nenne ich eine *Fredholmsche Gleichung*, wenn sie außer I bis III noch folgender Voraussetzung genügt:

IV. Es soll eine (ganze, positive, endliche) Zahl m vorhanden sein, so daß für jede Gleichung

$$(1.2) \quad W - \int K^{(n)} \cdot W' dt' = M_n \text{ in } b$$

im Fall $n \geq m$ alle Hauptsätze der Auflösungstheorie bestehen, die ich nachfolgend in Nr. 2 b zusammenstelle; dabei sei M_n von gleicher Art wie M . Das schließt nicht aus, daß dann die Hauptsätze auch bei einem solchen n zutreffen, welches kleiner als m ist.

c. Ferner nenne ich (1.1) eine *Levische Gleichung*, wenn sie I genügt und an Stelle von II bis IV folgende Bedingungen erfüllt:

V. $|K|$ sei integrierbar nach t und nach t' .

VI. $\int |K| dt$ und $\int |K| dt'$ seien gleichmäßig konvergent in b , das heißt, zu einer willkürlich vorgegebenen, positiven Zahl z lasse sich eine Zahl $g(z)$ stets so bestimmen, daß

$$\int |K(w, t')| dw < z \quad \text{für } R(t, w) \leq g(z)$$

und

$$\int |K(t, w)| dw < z \quad \text{für } R(w, t') \leq g(z)$$

gleichmäßig für t und t' in b gilt; dabei setze ich $R = R(t, t') \geq 0$ und

$$R^2 = \sum_v (x_v - x'_v)^2 \quad \text{für } v = 1 \dots q,$$

insbesondere

$$R^2 = (x - x')^2 + (y - y')^2 \quad \text{für } q = 2.$$

VII. K habe bei festem t eine für t in b beschränkte Anzahl von Unstetigkeitsstellen. Gleiches gelte in bezug auf t' .

Anmerkung. Bezeichne ich für ein bestimmtes t in b die dann noch für t' möglichen Unstetigkeitsstellen mit $t' = t'_1, t' = t'_2, \dots, t' = t'_n$, so ist $n = n(t)$. Es wird die Beschränktheit von $n(t)$ für t in b verlangt. — Der Zusammenhang von VI und VII mit III wird nicht erörtert.

VIII. Irgendein Glied von jeder der vier in I genannten Reihen sei beschränkt.

d. Alle bisher gemachten Annahmen ersetze ich durch die beiden folgenden:

IX. Es sei $K = N : R^p$ und $0 \leq p < q$ (q gleich der Anzahl der unabhängigen Veränderlichen). Hierin sei $N = N(t, t')$ stetig für t und t' in b , solange $R > 0$ ist, und im übrigen beschränkt.

X. M und W seien stetig in allen Punkten oder in allen Punkten mit Ausnahme einer endlichen Anzahl von Unstetigkeitsstellen. In letzteren sollen M und W beschränkt sein.

Dann liegt ein *gemeinsamer Sonderfall* der Fredholmischen und Levischen Gleichung vor (wie ich in Nr. 5b und 8c nachträglich und beiläufig beweise).

2. Hauptsätze der Auflösungstheorie für Integralgleichungen.

In folgender Nr. 2b stelle ich vier Hauptsätze für die Integralgleichung (1. 1) auf und untersuche anschließend, bei welchen Voraussetzungen für K , M und W diese Sätze alle oder teilweise gelten.

a. Dabei benutze ich die Begriffe *Defekt und Eigenlösungen*. Unter dem Defekt i der Integralgleichung (1. 1) verstehe ich die Anzahl der untereinander linear unabhängigen Lösungen der homogenen Gleichung, also von (1. 1) mit $M = 0$ (einschließlich dem Sonderfall einer einzigen, nicht identisch verschwindenden Lösung). Aus diesen erhalte ich im Fall $i > 0$ durch Orthogonalisierung und Normierung i Eigenlösungen $W_u = W_u(t)$ für $u = 1 \dots i$ (nach dem später in Nr. 4a angegebenen Verfahren, das ich hier auf reelle Funktionen zu spezialisieren habe). Entsprechend bezeichne ich im transponierten Fall (das heißt für K'' anstatt K) den Defekt mit i'' und die Eigenlösungen mit $W'_v = W'_v(t)$ für $v = 1 \dots i''$. Unterscheiden sich zwei Integralgleichungen durch ihre Kerne, so können doch ihre Eigenlösungen übereinstimmen (wie das später in Nr. 7c belegt wird).

b. Die Hauptsätze für (1. 1) sind:

I. *Erster Defektsatz*. i ist Null oder eine endliche (positive, ganze) Zahl.

II. *Lösungssatz*. Im Fall $i = 0$ hat die homogene Gleichung nur die Lösung Null; dann hat die inhomogene Gleichung für jedes M eine eindeutige Lösung W . Im Fall $i > 0$ ist die allgemeine Lösung der homogenen Gleichung gleich $\sum_u k_u \cdot W_u$ für $u = 1 \dots i$, wobei die k_u willkürliche Konstanten sind; dann ist die inhomogene Gleichung für solche und nur für solche M lösbar, die

$$(2. 1) \quad \int M \cdot W'_v dt = 0 \quad \text{für } v = 1 \dots i''$$

befriedigen. Ein solches M ist beispielsweise

$$M = M_0 - \sum_v W'_v \cdot \int M_0 \cdot W'_v dt,$$

wobei $M_0 = M_0(t)$ eine beliebige Funktion der Art wie M ist. i'' gibt die Anzahl der notwendigen und hinreichenden Lösungsbedingungen für (1. 1) bei allgemeinem M an; bei speziellem M kann diese Anzahl kleiner als i''

sein, wie das beispielsweise vor (4.11) der Fall sein wird. Ist W_0 irgendeine spezielle Lösung von (1.1), so ist demnach die allgemeine Lösung

$$(2.2) \quad W = W_0 + \sum_u k_u \cdot W_u \quad \text{für } u = 1 \dots i.$$

Im transponierten Fall gilt das Entsprechende.

III. *Zweiter Defektsatz.* $i = i''$.

IV. *Existenzsatz für den lösenden Kern.* Man nennt $L = L(t, t')$ einen lösenden Kern von (1.1), wenn

$$(2.3) \quad W = M - \int L \cdot M' dt'$$

eine Lösung von (1.1) ist; L soll für alle zulässigen M die gleiche Funktion sein; ferner erfülle L die allgemeinen Eigenschaften, die von K jeweils vorausgesetzt und festgestellt werden.

Satz: *Es gibt mindestens ein solches L zu (1.1), das ich mit $L_{00} = L_{00}(t, t')$ bezeichne.* — Man unterscheidet den „lösenden Kern im erweiterten Sinn“ für $i > 0$ (sogenannte Pseudoresolvente, vgl. Hell.-Toepl., S. 1374 und 1377) von dem „lösenden Kern im engeren Sinn“ für $i = 0$ (sogenannte Resolvente, vgl. Hell.-Toepl. und Courant).

Nach (2.2, 3) ist

$$(2.4) \quad W = M - \int L_{00} \cdot M' dt' + \sum_u k_u \cdot W_u$$

die allgemeine Lösung von (1.1), wenn M die dazu notwendigen und hinreichenden Bedingungen (2.1) erfüllt.

e. Vorstehende vier Hauptsätze sind bei stetigen und Schmidtschen K , ferner bei M und W , die der Voraussetzung I der Nr. 1 genügen, bekannt; die Stetigkeit darf an endlichvielen etwa $(q-1)$ -dimensionalen Ebenen unterbrochen werden, woselbst K endliche Sprünge machen soll. (Vgl. Hell.-Toepl., S. 1376–1377, 1385, 1388 = Nr. 12c unter Berücksichtigung von S. 1389 = Nr. 13a.) Der in der Literatur anzutreffende „Alternativsatz“ ist eine Zusammenstellung von erstem Defektsatz und Lösungssatz.

Die Hauptsätze können in vollem Umfang auch für allgemeinere Gleichungen (1.1) bestehen. Solche Gleichungen mögen im engeren Sinn uneigentlich singular heißen, zweckmäßigerweise rechne ich hierzu auch die Integralgleichungen mit stetigen K . Gilt der erste Defektsatz und der Lösungssatz, so nenne ich (1.1) im erweiterten Sinn uneigentlich singular. Dann sind weder $i = i''$ noch die Existenz von L_{00} bewiesen. Man nennt (1.1) eigentlich singular, wenn mindestens der Lösungssatz nicht in dem angegebenen Umfang gilt. (Hell.-Toepl., Nr. 12 und 21.) Die Definition der *Singularität einer Integralgleichung* und die gegenseitige Abgrenzung der möglichen Unterfälle wird in der Literatur nicht scharf umrissen, weil die Gesamtheit der im engeren Sinn und diejenige der im erweiterten Sinn uneigentlich singulären Gleichungen

unbekannt ist. Vor allem müßte man nach Hilbert die Voraussetzungen von K derart bestimmen, daß bei deren Erfüllung die Integralgleichung auf ein vollstetiges und allgemeiner auf ein gewisses beschränktes Gleichungssystem zurückgeführt und so der Theorie von Hilbert zugänglich gemacht wird. (Hell.-Toepl., Nr. 16, insbesondere Vorbemerkung und Absatz e, ferner Nr. 18 b 4 daselbst.) Weiter müßte man die Zurückführung auf noch umfassendere Theorien versuchen (vgl. Hell.-Toepl., S. 1416). Dabei bedenke man, daß „keine“ Darstellungsformeln für die Lösungen gewonnen werden, sondern nur ihre Existenz auf Grund eines den Häufungstellensatz benutzenden Auswahlverfahrens bewiesen wird“ (Hell. Nr. 7). (Die Andeutungen in Courant, Bd. 1, S. 133, ferner von Courant in den Annalen 89, S. 178, sind im Sinne von Hell.-Toepl., S. 1388 = Nr. 12c und S. 1397 = letzter Absatz von Nr. 15c zu verstehen.)

Die in Nr. 1 gemachten Voraussetzungen sind (worauf ich noch verschiedentlich zurückkomme) zur uneigentlichen Singularität hinreichend, aber nicht notwendig. Das gilt auch für die Hilbertsche Voraussetzung über K zur Umwandlung der Integralgleichung in ein lineares Gleichungssystem mit unendlichvielen Unbekannten, wonach K^2 in bezug auf t und auf t' integrierbar sein müßte. (Hell.-Toepl., Nr. 15, insbesondere letzter Absatz von c daselbst.) Zur Erläuterung ziehe ich einen eindimensionalen Ausgangsbereich b heran, etwa $0 \leq t \leq 1$. Ich setze

$$K = 1 : t^p \cdot t'^q \quad \text{für} \quad p \geq 1:2, \quad q > 0 \quad \text{und} \quad p + q < 1.$$

Weiter seien M und W beschränkt und integrierbar. Dann sind die Voraussetzungen III und VI der Nr. 1 nicht erfüllt. Ferner ist K^2 in bezug auf t nicht integrierbar. Jedoch ist aus

$$(2.5) \quad W - \int (W' : t^p \cdot t'^q) dt' = M$$

auf $\int (W : t^q) dt = -\frac{1-p-q}{p+q} \cdot \int (M : t^q) dt$, somit auf

$$(2.6) \quad W = M - \frac{1-p-q}{p+q} \cdot \frac{1}{t^p} \cdot \int (M : t^q) dt$$

zu schließen. Umgekehrt folgt aus (2.6)

$$\int (W' : t^p \cdot t'^q) dt' = -\frac{1-p-q}{p+q} \cdot \frac{1}{t^p} \cdot \int (M : t^q) dt = W - M,$$

somit (2.5). Also sind (2.5) und (2.6) einander gleichwertig. Folglich ist

$$i = 0 \quad \text{und} \quad L_{00} = (1-p-q) \cdot K : (p+q).$$

Ebenso ist im transponierten Fall $i'' = 0$ und der lösende Kern gleich $(1 - p - q) \cdot K'' : (p + q)$. Demnach bestehen die Hauptsätze für (2. 5) und das gilt für jede endliche Iteration von (2. 5), da

$$K^{(n)} = K : (1 - p - q)^{n-1}$$

ist.

d. Fredholm beweist einen Teil der Hauptsätze für (1. 1) mittels Umweg über (1. 2) (Literatur „Fredholm ...“). Er bestimmt $K = N : R^p$ etwas allgemeiner als in IX der Nr. 1, indem er N und auch M als endlich (worunter er „beschränkt“ versteht) und integrierbar annimmt. Es werden keine Voraussetzungen für W , keine für die Vertauschbarkeit von Integrationsfolgen und keine betreffs Unstetigkeitsstellen der vorkommenden Integranden erwähnt. Im einzelnen stellt Fredholm fest:

I. Die Fredholmsche Gleichung ist für $K = N$ (also $p = 0$) im engeren Sinn uneigentlich singular. (Fredholm, num. 8 u. 9; er verlangt von $K = N$ keine Stetigkeit, wie das Hell-Toepl., S. 1370 und 1386, und Courant, Bd. 1, S. 124, bei der Wiedergabe seiner Theorie der Einfachheit wegen annehmen.)

II. Die Hauptsätze gelten für (1. 2), wenn $n > q : (q - p)$ ist. Dann ist $K^{(n)}$ beschränkt (Fredh., num. 15 und 16). Die Vertauschbarkeit der Integrationsfolgen gilt auch in dem von Fredholm ausgeschlossenen Fall, daß $q : (q - p)$ eine ganze, positive Zahl ist. Denn dann ist

$$|K^{(v)}| < h \cdot (1 + |\log |x_1 - x'_1| \dots \log |x_q - x'_q||)$$

für $v = q : (q - p)$, h eine positive Konstante, also $K^{(v+1)}$ beschränkt und $n = v + 1 > q : (q - p)$. In der Voraussetzung IV der Nr. 1 wäre daher n gleich der größten in $(1 + \frac{q}{q-p})$ enthaltenen ganzen Zahl zu setzen.

Nachfolgend bezeichne ich den Defekt, die etwa vorhandenen Eigenlösungen und die allgemeine Lösung von (1. 1) ausführlicher als bisher mit $i(K)$, $W_u(K)$ und $W(K)$, ferner diejenigen von (1. 2) mit $i(K^{(n)})$, $W_u(K^{(n)})$ und $W(K^{(n)})$. Entsprechende Bezeichnungen gelten im transponierten Fall für K'' anstatt K .

III. $W(K)$ ist im Existenzfall in $W(K^{(n)})$ enthalten. (Fredh., num. 17, woselbst $q = 1$ gesetzt wurde.)

IV. Aus $i(K) = 0$ folgt $i(K'') = 0$ auch im Fall $i(K^{(n)}) > 0$ für $n \geq m$. (Fredh. num. 18), wenn

$$(2. 7) \quad \text{Determinante der } f_{uv} = \text{Determinante der } f''_{uv}$$

ist (Freihl. num. 14). Hierin sind f_{uv} und f''_{uv} gewisse Konstanten, für die folgende beiden Systeme von Gleichungen bestehen (auf die ich in Nr. 5c zurückkomme):

$$(2. 8) \quad W_u(K^{(n)}) - \frac{c}{1^n} \cdot \int K \cdot W_u(K^{(n)})' dt' = \sum_r f_{ur} \cdot W_r(K^{(n)}),$$

$u \text{ und } r = 1 \dots i(K^{(n)}),$

ferner

$$(2.9) \quad W_*(K^{(n)}) - 1^{\frac{c}{n}} \cdot \int K'' \cdot W_*(K^{(n'')})' dt' = \sum_{s,v} f_{s,v}' \cdot W_*(K^{(n'')}),$$

s und $v = 1 \dots i(K^{(n'')})$;

$1^{\frac{1}{n}}$ sei eine primitive n -te Einheitswurzel, etwa

$$(2.10) \quad 1^{\frac{1}{n}} = \cos \frac{2\pi}{n} + \sqrt{-1} \cdot \sin \frac{2\pi}{n},$$

und c eine der Zahlen $0, 1, \dots, (n-1)$. Fredholm setzt $c = 0$ und beweist (2.7) für den Fall, daß die Determinante der

$$\int W_u(K^{(n)}) \cdot W_v(K^{(n'')}) dt$$

von Null verschieden ist. Hierbei wird $i(K^{(n)}) = i(K^{(n'')}) = i(K^{(n'')})$ benutzt. „Un simple raisonnement par continuité permettant évidemment d'étendre la proposition au cas où le déterminant est nul“ (Fredh., S. 385). Das wird nicht bewiesen. Ich zeige später in Nr. 5c, daß (2.7) stets richtig ist. Dabei benutze ich jedoch die auf anderem Wege bewiesene Formel (4.5), welche die Fredholmsche Behauptung (aus $i = 0$ folgt $i'' = 0$) umfaßt. Iglisch schließt die Beweisücke selbst (wie nachfolgend in Nr. 2e angegeben wird). (Diesbezüglich ist Hell.-Toepl., Nr. 12a zu ergänzen.)

V. Wenn $W(K)$ existiert und $i(K) = i(K'') = 0$ gilt, so läßt sich $W(K)$ aus $W(K^{(n)})$ aussondern und $L = L(K)$ angeben. Der Fall $i = i'' = 0$ ist jedoch nicht von vornherein bekannt, wenn $i(K^{(n)}) > 0$ ist. Die Umkehrung, daß das durch L dargestellte W die ursprüngliche Integralgleichung löst und also ein W existiert, ferner der Fall $i > 0$ werden von Fredholm nicht behandelt. Das hole ich in Nr. 5a nach und zeige dort, daß die in Nr. 1b allgemeiner angegebene Fredholmsche Gleichung im engeren Sinn uneigentlich singular ist, ohne auf die Fredholmschen Ausführungen zurückzugreifen. (Das geschieht in determinantenfreier Weise; vgl. Nr. 10 in Hell.-Toepl.)

e. Iglisch beweist die beiden Defektsätze und den Lösungssatz für die „Fredholmsche Gleichung“ und benutzt dabei das Fredholmsche Iterationsverfahren in determinantenfreier Weise (vgl. Literatur „Iglisch . . .“; K'' heißt daselbst der adjungierte Kern). Er erwähnt keine Voraussetzungen für M und W . Das Kernstück seiner Theorie ist der Beweis dafür, daß $i'' = 0$ aus $i = 0$ folgt und umgekehrt (Igl., S. 228). Der Beweis läßt sich mit Hilfe der späteren Formel (4.4) wesentlich kürzen. Denn man kann im Fall $i(K) = 0$ durch passende Wahl von n stets den Fall $i(K^{(n)}) > 0$ überflüssig machen. Damit schließt Iglisch die Lücke bei Fredholm. Den Fall $i > 0$ führt er auf den Fall $i = 0$ zurück. Iglisch befaßt sich nicht mit dem Existenzsatz für den lösenden Kern. Doch folgt aus der Existenz von $W(K)$, daß das von Fredholm angegebene L im Fall $i = i'' = 0$ wirklich der lösende Kern ist.

1. Neben Fredholm hat Levi noch ein zweites Verfahren angegeben, um einen Teil der Hauptsätze für (1. 1) zu beweisen. Levi verlegt die Singularitäten des K der „Levischen Gleichung“ unter Anwendung des Kernabspaltungsverfahrens (Schmidt II) in den „kleinen Kern“. Dieser allein und nicht das ganze K wird zum Unterschied vom Fredholmschen Verfahren iteriert. (Vgl. Literatur „Levi . . .“. Levi gibt hinreichende Voraussetzungen an, mit denen die hier in Nr. 1c aufgestellten im wesentlichen übereinstimmen. Betreffs der hier in Nr. 1 VI angegebenen gleichmäßigen Konvergenz von $\int |K| dt$ und $\int |K| dt'$ weist Levi, S. 605, auf de la Vallée-Poussin hin: Sur la convergence des intégrales définies, Journal de Liouville, 1892, IV. serie, vol. 8, n. 61. Die beiden bei Levi angegebenen Bedingungen folgen auseinander. Also kann eine davon gestrichen werden.) Levi beweist den ersten Defektsatz und den Lösungssatz, somit die im erweiterten Sinn uneigentliche Singularität der Levischen Gleichung. Ferner zeigt er an Stelle von $i = i''$, daß aus $i'' = 0$ stets $i = 0$ und aus $i'' > 0$ stets $i > 0$ folgt. Die Existenz eines L kann er im allgemeinen nicht beweisen. Es ist noch nicht einmal die Existenz von $K^{(2)}$ bekannt (Levi, Anm. II der Nr. 3). Levi erreicht allgemeinere Aussagen bei Verzicht auf einen Teil der Voraussetzungen. Abschließend betont er, daß seine Theorie auf q -dimensionale Ausgangsbereiche b sinngemäß übertragbar ist.

Darüber hinaus zeige ich in Nr. 8a, daß auch für die Levische Gleichung $i = i''$ gilt. Ferner gebe ich in Nr. 8b den Zusammenhang zwischen $W(K)$ und $W(K'')$ an. Das bedeutet die Verallgemeinerung der nicht möglichen Darstellungen von $W(K)$ durch $L(K)$ und $W(K'')$ durch $L(K'')$ und ihres Zusammenhanges mittels $L(K'') = L(K)''$. Aus diesem Zusammenhang folgere ich in Nr. 8c auf den noch ausstehenden Existenzsatz für den lösenden Kern bei Spezialisierung auf den gemeinsamen Sonderfall der Fredholmschen und Levischen Gleichungen, für den $K = N: R^p$ angenommen wurde. Außer diesen und den am Schluß von Nr. 2d genannten Vervollständigungen der Beweise der Hauptsätze der Auflösungstheorie für die Fredholmschen und Levischen Gleichungen werden über die Literatur hinausgehend die allgemeinsten Eigenschaften des lösenden Kernes L zu der Gleichung (1. 1) unter den Voraussetzungen I bis III der Nr. 1 in Nr. 6 festgestellt und in Nr. 7 für Sonderfälle ergänzt.

3. Integralabschätzungen als Hilfsformeln.

Ich stelle die zum restlichen Beweis der Hauptsätze und gelegentlich bei der P -Methode benutzten Integralabschätzungen unter Beschränkung auf den zweidimensionalen Fall zusammen. Man kann die Abschätzungsverfahren auf den q -dimensionalen Fall übertragen. Es lassen sich auch sonstige Verall-

gemeinerungen durchführen, beispielsweise betreffs Stetigkeitsverhältnisse der einführenden Funktionen.

Ich schreibe $b = b_0$, wenn

1. für jedes $t = (x, y)$ in einer gewissen Umgebung des ganzen Randes d von b ein eindeutiger Normalenabstand $n = n(t)$ des Punktes t von d vorhanden ist,

2. eine positive Zahl m existiert, so daß die Bedingung 1 für alle t mit einem $n(t)$ in dem System der Randstreifen $0 \leq n(t) \leq m$ gilt,

3. der Fußpunkt der Normalen durch t auf d eine eindeutige Bogenlänge s hat. (Hierauf gehe ich in Nr. 9 zwecks späteren Anwendungen ausführlich ein.)

In den Hilfsformeln beschränke ich mich nur dann auf $b = b_0$, wenn im Integranden eine durch $n = n(t)$ voraussetzungsgemäß abgeschätzte Funktion auftritt.

Die Hilfsformeln gehen schrittweise auseinander hervor. Die Beweise fußen auf geeigneten Zerlegungen der Integrationsbereiche und der Integranden bei gelegentlicher Verwendung von passenden Integrationsveränderlichen. Hierbei führe ich folgende, in der weiteren Arbeit ständig wiederkehrende Zeichen ein. Für zwei Ausdrücke g und h , die Konstanten oder Funktionen bedeuten, sei $|g| = |k \cdot h|$, wobei k eine positive, endliche Konstante oder eine beschränkte Funktion größer als eine positive Konstante und im übrigen unwesentlich ist. Dann schreibe ich $g \equiv h$. Entsprechend setze ich $\lesseqgtr, \gtrless, \gtrless, \lesseqgtr$ an Stelle von $\leq, \geq, >, <$, wenn diesbezüglich $<, \leq, >, \geq$ an Stelle von $=$ steht. Es gilt \lesseqgtr , wenn gleichzeitig \gtrless und \lesseqgtr ist. Diese Zeichen wende ich auch bei fortlaufenden Abschätzungen an.

a. Zuerst stelle ich die *Kleinheit gewisser Integrale* fest. $R = R(t, t')$ sei (wie immer in dieser Arbeit) die Entfernung zwischen t und t' ; $R \geq 0$. p bzw. a sei irgendeine Konstante, die $0 \leq p < 2$ bzw. $0 \leq a < 1$ genügt. Dann ist

$$\int \frac{1}{R^p} dt = \frac{2\pi}{2-p} \cdot a^{2-p},$$

wenn dt über die Kreisfläche um t' als Mittelpunkt mit dem Halbmesser a integriert wird. Diese Formel wird nachfolgend verallgemeinert. Hierzu werde dt über einen Teilbereich von b integriert, der von gleicher Art wie b sei und den Inhalt $\pi \cdot a^2$ habe; als solchen Teilbereich wähle ich etwa irgendeine zu b gehörende Kreisfläche A . Ferner führe ich eine Funktion $f = f(t)$ folgendermaßen ein. Ist b nicht von der Art b_0 , so sei f irgendeine in b integrierbare und beschränkte Funktion. Ist aber b von der Art b_0 , so verstehe ich unter e eine Konstante, die $2e < 2 - p$ und $0 \leq e < 1$ genügt und sonst willkürlich ist; dann sei f irgendeine in b integrierbare und in $(b - d)$ be-

schränkte Funktion, die $f \lesssim 1 : n^\epsilon$ für $0 \leq n(t) \leq m$ erfüllt. Insbesondere kann $f \lesssim \log n$ für $0 \leq n(t) \leq m$ sein. Denn es gilt stets

$$(3.4) \quad \log n \lesssim 1 : (v \cdot u^\epsilon)$$

für irgend zwei positive Zahlen u und v . Schließlich sei z irgendeine Konstante, die $0 < z < 2 \cdot (1 - \epsilon)$ genügt.

Ich behaupte

$$(3.1) \quad \int_A (f : R^p) dt \lesssim a^{2-p-2\epsilon},$$

$$(3.2) \quad \int f \cdot \log R dt \lesssim a^{2-z-2\epsilon}$$

und

$$(3.3) \quad \int f dt \lesssim a^{2-2\epsilon},$$

für beschränkte f gelten diese Formeln mit $\epsilon = 0$.

Zum Beweis bezeichne ich mit A' die Kreisfläche um t' mit dem Halbmesser a und mit B den etwa vorhandenen gemeinsamen Teil von A' und A . Dann gilt

$$\int_A (f : R^p) dt \lesssim \int_{A-B} (|f| : R^p) dt + \int_{A'} (|f| : R^p) dt,$$

also, wenn man f als beschränkt annimmt,

$$\lesssim (\pi \cdot a^2 : a^p) + (2 \cdot \pi \cdot a^{2-p} : (2-p)) \simeq a^{2-p}$$

für $a > 0$, woraus (3.1) mit $\epsilon = 0$ folgt. Bei Einschränkung von b auf b_0 wird $\int (f : R^p) dt$ ebenso durch a^{2-p} und damit erst recht durch $a^{2-p-2\epsilon}$ abgeschätzt, solange dt über A außerhalb des Randstreifens integriert wird. Ich beschränke daher im folgenden A , A' und B auf den Randstreifen. Jetzt ist

$$\int_{A-B} \frac{|f|}{R^p} dt \lesssim \frac{1}{a^p} \cdot \iint \frac{dn ds}{n^\epsilon} \lesssim \frac{a^{2-p-\epsilon}}{1-\epsilon},$$

da n und s jeweils über ein Intervall von geringerer Länge als $2 \cdot a$ zu integrieren sind. Weiter ist bei Integration über denjenigen Teil von A' , für den $n \geq R^2$ gilt,

$$\int \frac{|f|}{R^p} dt \lesssim \int \frac{dt}{R^{p+2\epsilon}} \lesssim \frac{a^{2-p-2\epsilon}}{2-p-2\epsilon}.$$

Ferner ist bei Integration über den restlichen Teil von A' , für den $n < R^2$ gilt,

$$\int \frac{|f|}{R^p} dt \lesssim \iint \frac{dn ds}{n^{\frac{p}{2}+\epsilon}} \lesssim \frac{a^{2-\frac{p}{2}-\epsilon}}{1-\frac{p}{2}-\epsilon}.$$

Aus allem folgt (3. 1) für b_0 . Aus (3. 1) für b und b_0 ergibt sich (3. 2), wenn man (3. 4) berücksichtigt. Schließlich ist (3. 3) die Formel (3. 1) für $p = 0$.

b. Weiter untersuche ich die *Stetigkeit und Singularität gewisser Integrale*. Ich schicke voraus, daß bei letzteren dt über einen Teilbereich B von b integriert wird, der von gleicher Art wie b sein möge und auch mit b zusammenfallen kann. w sei ein zweiter Parameterpunkt neben t' . p und q seien irgendwelche Konstanten, die $0 \leq p < 2$ und $0 \leq q < 2$ genügen. f habe die in Nr. 3a angegebenen Eigenschaften, wobei neben $2e < 2 - p$ noch $3e < 2 - q$ bestehe. Weiter seien $g = g(t, t')$, $h = h(t, t')$ und $N = N(t, t')$ irgendwelche in b für $R > 0$ stetige und überall in b beschränkte Funktionen. Schließlich sei $J = J(t, t')$ irgendeine in b stetige Funktion. Dann gelten

$$\int \frac{1}{R(t, w)} \cdot \frac{1}{R(w, t')} dw = N \cdot \log R$$

und

$$\int g(t, w) \cdot \log R(t, w) \cdot \frac{1}{R(w, t')} \cdot f(w) dw = J$$

für beschränkte f (Lichtenstein, konf. Abbildungen, Formel 57 und 58 auf S. 24, ferner S. 28 unten).

Allgemeiner behaupte ich

$$(3.6) \quad \int \frac{g(t, w)}{R^p(t, w)} \cdot \frac{h(w, t')}{R^q(w, t')} \cdot f(w) dw \\ = \frac{N}{R^{p+q-2+2e}} \quad \text{für } p+q-2+2e > 0, \\ \text{oder} = N \cdot \log R \quad \text{für } p+q-2+2e = 0, \\ \text{oder} = J \quad \text{für } p+q-2+2e < 0;$$

für beschränkte f ist $e = 0$ zu setzen.

Ferner besteht

$$(3.7) \quad \int g(t, w) \cdot \log R(t, w) \cdot \frac{h(w, t')}{R^q(w, t')} \cdot f(w) dw = J, \\ \int g(t, w) \cdot \frac{h(w, t')}{R^q(w, t')} \cdot f(w) dw = J, \\ \int g(t, w) \cdot \log R(t, w) \cdot \log R(w, t') \cdot h(w, t') \cdot f(w) dw = J, \\ \int g(t, w) \cdot \log R(t, w) \cdot h(w, t') \cdot f(w) dw = J$$

und Entsprechendes bei Vertauschung von t mit t' . Zuvor beweise ich noch, daß

$$(3.5) \quad \int \frac{f \cdot g}{R^p} dt \quad \text{stetig}$$

ist.

Zum Beweise von (3.5) gebe ich irgend zwei Punkte u und t' im Innern einer festen Kreisfläche A mit dem Halbmesser a an, wobei ich von A den nicht zu B gehörenden Teil weglasse. Dann ist die Differenz

$$\int_B \frac{f \cdot g(t, t')}{R^p(t, t')} dt - \int_B \frac{f \cdot g(t, u)}{R^p(t, u)} dt \\ \lesssim \int_{B-A} \left| \frac{f \cdot g(t, t')}{R^p(t, t')} - \frac{f \cdot g(t, u)}{R^p(t, u)} \right| dt + \int_A \frac{|f| dt}{R^p(t, t')} + \int_A \frac{|f| dt}{R^p(t, u)}.$$

Weiter ist gemäß (3.1) die Differenz $\lesssim a^{2-p-2\epsilon}$ für $\lim u = t'$ und das für jedes a , woraus (3.5) folgt. Nach gleichem Verfahren mit $(p+q)$ anstatt p erhält man (3.6) für $2 - (p+q) - 2\epsilon > 0$. Somit ist in den weiteren Beweisteilen der Fall $p+q-2+2\epsilon < 0$ als erledigt anzusehen. Ferner ist zu berücksichtigen, daß (3.6) aus (3.5) folgt, solange $R > 0$ ist. Man integriere über einen Bereich, der t im Innern enthält und t' ausschließt, etwa eine Kreisfläche um t mit dem Halbmesser $R:2$, von der ich den nicht zu B gehörenden Teil weglasse. Dann ist das Integral in (3.6) nach (3.1) abzuschätzen durch

$$\lesssim \frac{1}{R^q} \cdot \int \frac{|f(w)|}{R^p(t, w)} dw \lesssim \frac{1}{R^{p+q-2+2\epsilon}}$$

für

$$p+q-2+2\epsilon > 0$$

oder $\lesssim 1$ für

$$p+q-2+2\epsilon = 0.$$

Weiter integriere man über den durch $R:2 \leq R(t, w) \leq R(w, t')$ für w in B gekennzeichneten Bereich und unterscheide:

- (I) w liegt außerhalb des Randstreifens,
- (II) w liegt im Randstreifen so, daß $n(w) \geq R^2$ ist,
- (III) $n(w) < R^2$.

Dann ist das Integral (3.6) abzuschätzen durch

$$\int \frac{|f(w)|}{R^{p+q}(t, w)} dw$$

und somit

- (I) Integral $\lesssim \frac{1}{R^{p+q-2}}$ für $p+q > 2$ oder $\lesssim \log R$ für $p+q = 2$,
- (II) Integral $\lesssim \frac{1}{R^{2\epsilon}} \cdot \int \frac{dw}{R^{p+q}(t, w)} \lesssim \frac{1}{R^{p+q-2+2\epsilon}}$ für $p+q > 2$
oder $\lesssim \frac{1}{R^{2\epsilon}} \cdot \log R$ für $p+q = 2$,
- (III) Integral $\lesssim \frac{1}{R^{p+q}} \cdot \int_0^{R^2} \frac{dn}{n^{\epsilon}} = \frac{1}{(1-\epsilon) \cdot R^{p+q-2+2\epsilon}}$
für $p+q-2+2\epsilon \geq 0$.

Bei Vertauschung von t mit t' erhält man die gleichen Abschätzungsergebnisse. Aus allem folgt (3. 6) vollständig. Schließlich ergibt sich (3. 7) aus (3. 6) mit Hilfe von (3. 4). Man braucht in (3. 4) mit $u = R$ nur $v < 2 - p - 2e$, $v < 2 - q - 2e$ und $v < 1 - e$ zu wählen, was mit $v > 0$ verträglich ist.

e. Ferner gelten nachstehende Darstellungen der *Iterationen* von $K = N : R^p$. Dabei sollen N und J die in Nr. 3b angegebenen Eigenschaften haben. p sei irgendeine Konstante, die $0 < p < 2$ genügt. $K^{(v)}$ sei der $(v - 1)$ -te iterierte Kern. Weiter sei m die durch

$$\frac{2}{2-p} < m \leq 1 + \frac{2}{2-p}$$

eindeutig bestimmte ganze Zahl; also ist

$$m = m(p) \geq 2$$

und

$$\frac{2m-4}{m-1} \leq p < \frac{2m-2}{m}.$$

Dann besteht:

$$K^{(v)} = N : R^{2-v(2-p)}$$

für $v = 1 \dots (m-2)$, wenn $m > 2$ ist.

$$K^{(m-1)} = N : R^{2-(m-1)(2-p)}$$

für $m < 1 + \frac{2}{2-p}$, das heißt $\frac{2}{2-p}$ nicht ganzzahlig, oder

$$K^{(m-1)} = N \cdot \log R$$

für $m = 1 + \frac{2}{2-p}$, das heißt $\frac{2}{2-p}$ eine der Zahlen 2, 3, ...

$$K^{(v)} = J$$

für $v = m, m+1, \dots$, also für $v \geq 2$, wenn $0 < p < 1$,

ferner für $v \geq 3$, wenn $1 \leq p < 4:3$ usw.

Diese Formeln ergeben sich aus der wiederholten Anwendung von (3. 6) und (3. 7), wenn man die verschiedenen Möglichkeiten durchrechnet.

d. Schließlich zeige ich die *Vertauschbarkeit gewisser Integrationsfolgen*. Multipliziert man das in (3. 6) angeschriebene Integral mit $j = j(t)$, wobei j von gleicher Art wie f sein soll, und integriert dt und dw über b , so behaupte ich die Vertauschbarkeit der Integrationsfolgen, also

$$(3. 9) \quad \iint \frac{g(t, w)}{R^p(t, w)} \cdot \frac{h(w, t')}{R'(w, t')} \cdot f(w) \cdot j(t) dw dt = \iint \dots dt dw.$$

Das fußt auf der ferner behaupteten Formel

$$(3.8) \quad \iint \frac{N \cdot f \cdot j}{R^p} dt dt' = \iint \frac{N \cdot f \cdot j}{R^p} dt' dt,$$

wobei dt über B und dt' über B' integriert werden; B' sei von gleicher Art wie B (vgl. Nr. 3b); B und B' dürfen auch zusammenfallen.

Zum Beweis von (3.8) führe ich $Z = Z(t, t')$ für t in B und t' in B' folgendermaßen ein. Es sei $Z = 0$, wenn mindestens einer der drei nachstehenden Fälle eintritt:

$$R \leq a, \quad n \leq a^2, \quad n' \leq a^2.$$

Ferner sei $Z = N$, wenn $R \geq 2 \cdot a$ und dazu gegebenenfalls im Randstreifen $n \geq 2 \cdot a^2$ und $n' \geq 2 \cdot a^2$ ist. Hierin bedeute a eine positive Zahl. Im übrigen sei Z eine überall stetige Funktion. (3.8) gilt für Z anstatt N . Also ist die Differenz der beiden in (3.8) angeführten Doppelintegrale gleich

$$\iint \frac{f \cdot j \cdot (N - Z)}{R^p} dt dt' - \iint \dots dt' dt.$$

Nach (3.1, 3) gibt es eine positive Zahl c derart, daß beide Doppelintegrale $< a^e$ für alle a sind. Also ist die genannte Differenz gleich Null, das heißt, es besteht (3.8).

Zum Beweis von (3.9) bezeichne ich mit A bzw. C den in b gelegenen Teil einer Kreisfläche um t' mit dem Halbmesser a bzw. c und setze $a < c$. Dann ist die Differenz der beiden in (3.9) angeführten Doppelintegrale gleich

$$\begin{aligned} & \int_{b-A} \int_b \dots dt dw + \int_A \int_b \dots dt dw - \int_b \int_{b-A} \dots dw dt - \\ & - \int_{b-C} \int_A \dots dw dt - \int_C \int_A \dots dw dt. \end{aligned}$$

Hierin hebt sich nach (3.8) das erste Doppelintegral gegen das dritte weg.

Im zweiten Doppelintegral ist $\int \frac{g(t, w)}{R^p(t, w)} \cdot j(t) dt$ nach (3.5) stetig, also das zweite Doppelintegral nach (3.1) durch $> a^{2-q-2e}$ abzuschätzen. Im vierten Doppelintegral ist $R(t, w) \geq c - a$, also das vierte Doppelintegral nach (3.1) durch $< a^{2-q-2e} : (c - a)^p$ abzuschätzen. Im fünften Doppelintegral hat $\int \dots dw$ bis auf den Faktor $j(t)$ die in (3.6) angegebenen Werte, wobei ich auf eine Abschätzung mittels a verzichte. Also ist das fünfte Doppelintegral nach (3.1) durch $< c^{4-p-q-5e}$ für $p + q - 2 + 2 \cdot e > 0$ oder nach (3.2) durch $< c^{2-1-2e}$ für $p + q - 2 + 2e = 0$ oder nach (3.3) durch $< c^{2-2e}$ für $p + q - 2 + 2e < 0$ abzuschätzen. Das gilt für alle erlaubten a und c . Ich lasse zuerst a und danach c gegen Null streben. Dann ergibt sich, daß die Differenz der beiden in (3.9) angeführten Doppelintegrale gleich Null, das heißt, daß (3.9) richtig ist.

4. Zusammenhang der ursprünglichen mit der iterierten Integralgleichung.

Der Beweis der Hauptsätze mittels Iteration der ganzen Integralgleichung fußt auf folgendem Zusammenhang der ursprünglichen Integralgleichung (1. 1) mit der iterierten Gleichung (1. 2).

a. Ich bezeichne den Defekt, die etwa vorhandenen *Eigenlösungen* und die allgemeine Lösung von (1. 2) mit $i(K^{(n)})$, $W_e(K^{(n)})$ und $W(K^{(n)})$. Entsprechend bezeichne ich diejenigen von

$$(4. 1) \quad W - \frac{c}{1^n} \cdot \int K \cdot W' dt' = M$$

mit $i(1^n \cdot K)$, $W_u(1^n \cdot K)$ und $W(1^n \cdot K)$ für $c = 0 \dots (n-1)$; 1^n wie in (2. 10). Ferner ersetze ich im transponierten Fall überall K durch $K'' = K(t', t)$ und beachte $K^{(n)''} = K^{(n)}$. In dieser Nr. 4 genügt es, die Eigenlösungen als irgendein System von untereinander linear unabhängigen Lösungen der homogenen Gleichung anzusehen und zum Unterschied von Nr. 2a auf die Normierung und Orthogonalisierung zu verzichten. Die $W_u(1^n \cdot K)$ sind komplexe Funktionen der reellen Veränderlichen (x, y) .

Man kann jedoch ein System von irgendwelchen komplexen Funktionen $Z_u = Z_u(t)$ der reellen Veränderlichen (x, y) , von denen das Produkt je zweier integrierbar ist, in das folgende sogenannte unitäre System überführen. (Hell-Toepl., S. 1536; nachfolgendes ist eine Abänderung der Hilfsformel 4 auf S. 1394 daselbst.) Man setze

$$Y_1 = Z_1 : \left(\int Z_1 \cdot \bar{Z}_1 dt \right)^{\frac{1}{2}},$$

wobei die Tilde den Übergang zur konjugiert komplexen Funktion bedeutet, ferner

$$Y_v = (Z_v - \sum_u Y_u \cdot \int Z_v \cdot \bar{Y}_u dt) : \left(\int \text{Zähler} \times \overline{\text{Zähler}} dt \right)^{\frac{1}{2}}$$

für $u = 1 \dots (v-1)$ und $v = 2, 3, \dots$,

wenn der Zähler nicht identisch verschwindet, und

$$Y_v = 0,$$

wenn der Zähler identisch verschwindet. Die aufeinanderfolgenden, nicht identisch verschwindenden Y_v , von denen einige Indizes ausfallen können, sind ein System mit der Normierung

$$\int Y_u \cdot \bar{Y}_u dt = 1$$

und der Orthogonalisierung

$$\int Y_u \cdot \bar{Y}_v dt = 0 \quad \text{für } u \neq v,$$

das als unitär bezeichnet wird.

b. K und W sollen den Voraussetzungen I und II der Nr. 1 genügen. Aus den Eigenschaften der primitiven n -ten Einheitswurzel (2. 10) folgt:

Satz I. Setzt man die n linearen homogenen Integraltransformationen von W , nämlich

$$W - \frac{v}{1^n} \cdot \int K \cdot W' dt' \quad \text{für } v = 0 \dots (n-1),$$

in irgendeiner Reihenfolge zusammen, so erhält man

$$W - \int K^{(n)} \cdot W' dt'.$$

Satz II. Läßt man dabei die Transformation für $v = c$ aus, so erhält man

$$T_{n,c}(W) = W + \sum_e \frac{c}{1^n} \cdot \int K^{(e)} \cdot W' dt' \quad \text{für } e = 1 \dots (n-1),$$

woraus

$$(4.2) \quad \sum_e T_{n,e}(W) = n \cdot W \quad \text{für } c = 0 \dots (n-1)$$

folgt.

c. K , M und W sollen den Voraussetzungen I, II und IV der Nr. 1 genügen. Danach gibt es eine Zahl m , so daß für (1. 2) im Fall $n \geq m$ alle Hauptsätze gelten. Ich beweise

Satz III. Die $W_u(\frac{c}{1^n} \cdot K)$ stellen für $u = 1 \dots i(\frac{c}{1^n} \cdot K)$ und $c = 0 \dots (n-1)$ ein vollständiges System von untereinander linear unabhängigen Lösungen von (1. 2) mit $M_n = 0$ dar, soweit $i(\frac{c}{1^n} \cdot K) > 0$ besteht; es ist

$$(4.3) \quad \sum_e i(\frac{c}{1^n} \cdot K) = i(K^{(n)}) \quad \text{für } c = 0 \dots (n-1) \text{ und } n \geq m.$$

Die nicht identisch verschwindenden $W_u(\frac{c}{1^n} \cdot K)$ können daher bei Verzicht auf Normierung und Orthogonalisierung als $W_e(K^{(n)})$ für $v = 1 \dots i(K^{(n)})$ benutzt werden.

Aus Satz I folgt die bekannte Ungleichung $i(\frac{c}{1^n} \cdot K) \leq i(K^{(n)})$, wenn c irgendeine der Zahlen $0 \dots (n-1)$ ist. (Hell-Toepl., Nr. 11b, Satz anschließend (4), ferner Schmidt I, S. 448 oben.) Also sind alle $i(\frac{c}{1^n} \cdot K)$ endlich. Man kann daher

$$\sum_e \sum_u h_{eu} \cdot W_u(\frac{c}{1^n} \cdot K) = 0$$

für $u = 1 \dots i(1^{\frac{e}{n}} \cdot K)$ und $c = 0 \dots (n-1)$ mit den zu bestimmenden Zahlen h_{eu} ansetzen, soweit $i(1^{\frac{e}{n}} \cdot K) > 0$ besteht. Aus $T_{ne}(\Sigma_e \Sigma_u \dots) = 0$ folgt (nach Satz I) $T_{ne}(\Sigma_u h_{eu} \cdot W_u(1^{\frac{e}{n}} \cdot K)) = 0$. Wegen

$$W_u(1^{\frac{e}{n}} \cdot K) - 1^{\frac{v}{n}} \cdot \int K \cdot W_u(1^{\frac{e}{n}} \cdot K)' dt' = (1 - 1^{\frac{v-c}{n}}) \cdot W_u(1^{\frac{e}{n}} \cdot K)$$

ist daher

$$\Sigma_u h_{eu} \cdot W_u(1^{\frac{e}{n}} \cdot K) = 0,$$

somit

$$h_{eu} = 0 \quad \text{für } u = 1 \dots i(1^{\frac{e}{n}} \cdot K) \quad \text{und } c = 0 \dots (n-1).$$

Demnach bilden die $W_u(1^{\frac{e}{n}} \cdot K)$ ein System von $\Sigma_e i(1^{\frac{e}{n}} \cdot K)$ untereinander linear unabhängigen Funktionen. Da jedes $W_u(1^{\frac{e}{n}} \cdot K)$ von den $W_v(K^{(n)})$ für $v = 1 \dots i(K^{(n)})$ linear abhängt, ist $\Sigma_e i(1^{\frac{e}{n}} \cdot K) \leq i(K^{(n)})$. Hierbei benutze ich den aus der Gleichungstheorie stammenden

Hilfssatz IV. Sind in einem Gleichungssystem der Art

$$F_u(t) = \Sigma_v g_{uv} \cdot C_v(t) \quad \text{für } v = 1 \dots i \text{ und } u = 1 \dots j$$

die F_u untereinander linear unabhängig, so ist j kleiner oder gleich der kleinstmöglichen Zahl i ; letztere ist die festbleibende Anzahl von linear unabhängigen Funktionen unter den C_v . Die Funktionen und Zahlen können auch komplex sein.

Andererseits hängt $T_{ne}(W_v(K^{(n)}))$ nach Satz II von den $W_u(1^{\frac{e}{n}} \cdot K)$ für $u = 1 \dots i(1^{\frac{e}{n}} \cdot K)$ linear ab. Dann folgt aus (4. 2), daß $n \cdot W_v(K^{(n)})$ von den $W_u(1^{\frac{e}{n}} \cdot K)$ für $u = 1 \dots i(1^{\frac{e}{n}} \cdot K)$ und $c = 0 \dots (n-1)$ linear abhängt; $v = 1 \dots i(K^{(n)})$. Somit ist nach Hilfssatz IV $i(K^{(n)}) \leq \Sigma_e i(1^{\frac{e}{n}} \cdot K)$. Aus allem folgt Satz III.

Diesen kann man als Verallgemeinerung eines Satzes von Schmidt auffassen: Man bestimme

$$F_{ev} = F_{ev}(t) = W_v(K^{(n)}) + \Sigma_e 1^{\frac{e}{n}} \cdot \int K^{(e)} \cdot W_v(K^{(n)})' dt'$$

$$\text{für } e = 1 \dots (n-1), \quad c = 0 \dots (n-1) \quad \text{und } v = 1 \dots i(K^{(n)});$$

dann gibt es gewisse Konstanten g_{evu} , so daß

$$F_{ev} = \Sigma_u g_{evu} \cdot W_u(1^{\frac{e}{n}} \cdot K) \quad \text{für } u = 1 \dots i(1^{\frac{e}{n}} \cdot K)$$

besteht; es ist

$$n \cdot W_n(K^{(n)}) = \sum_r F_{r,n}.$$

(Vgl. Schmidt I, S. 447–448; die Untersuchung daselbst ist von der Definition der iterierten Kerne an bis einschließlich den beiden Sätzen anschließend Formel (2) ebenso für unsymmetrische Kerne gültig, die den Voraussetzungen I und II meiner Nr. 1 genügen; die weitere Untersuchung von Schmidt ist jedoch nur für symmetrische K möglich.)

Dieser Satz folgt aus Satz II. Ein bekannter Sonderfall von (4.3) ist:

$$\text{aus } i(K^{(n)}) \geq 1 \text{ folgt } \sum_c i(1^{\frac{c}{n}} \cdot K) \geq 1 \text{ für } c = 0 \dots (n-1).$$

(Hell-Toepl., Schlußsatz der Nr. 11b daselbst.)

d. Unter den Voraussetzungen I, II und IV der Nr. 1 beweise ich

Satz V. *Durchläuft n die Primzahlen $\geq m$, so gibt es höchstens endlich viele Werte n , für die nicht alle $i(1^{\frac{c}{n}} \cdot K)$ verschwinden; $c = 1 \dots (n-1)$; folglich gibt es unendlichviele Werte $n = g \geq m$, für die*

$$i(1^{\frac{c}{g}} \cdot K) = 0, \quad c = 1 \dots (g-1),$$

und damit

$$(4.4) \quad i(K^{(g)}) = i(K)$$

gilt. — Man kann demnach erreichen, daß der Defekt der iterierten Gleichung gleich dem der ursprünglichen ist.

Für $i(K^{(n)}) = 0$ sind nach (4.3) alle $i(1^{\frac{c}{n}} \cdot K) = 0$. Für $i(K^{(n)}) \geq 1$ gibt es nach (4.3) mindestens ein $v = v(n)$, so daß $1^{\frac{v}{n}}$ ein Eigenwert von K und damit $1^{\frac{v}{n} \cdot m}$ einer von $K^{(m)}$ ist; $0 \leq v \leq n-1$. $1^{\frac{v}{n} \cdot m}$ nimmt im Fall $v(n) \geq 1$ ständig neue Werte an, wenn n die Primzahlen $\geq m$ durchläuft. Da $K^{(m)}$ auf dem Einheitskreis der komplexen Zahlenebene höchstens endlichviele Eigenwerte besitzt, gilt $v(n) = 0$, wenn man von höchstens endlichvielen Werten n absieht. Hierbei habe ich eine Schlußweise von Iglisch (S. 228 daselbst) verwandt. Der Beweis von Iglisch läßt sich mit (4.4) wesentlich kürzen, wie in Nr. 2e erwähnt wurde.

e. Setzt man wieder I, II, IV der Nr. 1 voraus und bedeuten r und s irgendwelche ganze, positive Zahlen, so gilt der zweite Defektsatz:

$$(4.5) \quad i(1^{\frac{r}{s}} \cdot K) = i(1^{\frac{r}{s}} \cdot K'').$$

Denn (4.4) gilt auch im transponierten Fall. Also gibt es unendlichviele Werte $n \geq m$, für die gleichzeitig $i(K^{(n)}) = i(K)$ und $i(K^{(n)'}) = i(K'')$ ist,

woraus $i(K) = i(K'')$ folgt. Daran ändert sich nichts, wenn man von vornherein K durch $1^{\frac{r}{n}} \cdot K$ ersetzt und unabhängig davon mit den bisherigen Faktoren $1^{\frac{e}{n}}$ umgeht.

1. Die lösenden Kerne L von (1. 2) und (4. 1) werden mittels (2. 3) bestimmt, wobei L unabhängig von M_n und M ist. Insbesondere bezeichne ich L für (1. 2) mit $L(K^{(n)})$ und L für (4. 1) mit $L(1^{\frac{e}{n}} \cdot K)$. Ich beweise unter den Voraussetzungen I, II und IV der Nr. 1 die Richtigkeit von

$$(4. 6) \quad L(1^{\frac{e}{n}} \cdot K) = - \sum_e 1^{\frac{e}{n}} \cdot K^{(e)} + L(K^{(n)}) \\ + \sum_e 1^{\frac{e}{n}} \cdot \int L(K^{(n)}; t, z) \cdot K^{(e)}(z, t') dz \\ + \sum_v \sum_u W_u(1^{\frac{v}{n}} \cdot K) \cdot H'_{cuv}$$

für $e = 1 \dots (n-1)$, $u = 1 \dots i(1^{\frac{r}{n}} \cdot K)$ und $v = 0 \dots (n-1)$ außer $v = c$, c eine der Zahlen $0 \dots (n-1)$, soweit $i(1^{\frac{r}{n}} \cdot K) > 0$ besteht; hierin sind die $H_{cuv} = H_{cuv}(t)$ gewisse durch $L(K^{(n)})$ bestimmte und von M unabhängige Funktionen mit den gleichen allgemeinen Eigenschaften wie etwa $\int M \cdot K^{(n)} dt$.

Nach Satz II ist

$$(4. 7) \quad W(1^{\frac{e}{n}} \cdot K) = T_{nc}(W(K^{(n)})) \text{ für } M_n = M$$

eine Lösung von (4. 1), wenn c irgendeine der Zahlen $0 \dots (n-1)$ ist. Die notwendigen und hinreichenden Bedingungen zur Existenz von $W(K^{(n)})$ für $M_n = M$ und damit von $W(1^{\frac{e}{n}} \cdot K)$ in der Form (4. 7) sind nach (2. 1)

$$\int M \cdot W_v(K^{(n)'}) dt = 0 \quad \text{für } v = 1 \dots i(K^{(n)'})$$

wenn $i(K^{(n)'}) > 0$ ist. Sie gehen nach Satz III über in

$$(4. 8) \quad \int M \cdot W_u(1^{\frac{v}{n}} \cdot K'') dt = 0 \quad \text{für } u = 1 \dots i(1^{\frac{v}{n}} \cdot K'')$$

und $v = 0 \dots (n-1)$, soweit $i(1^{\frac{v}{n}} \cdot K'') > 0$ besteht, wenn $i(K^{(n)}) > 0$ ist.

Hiervon sind, wie sich durch Multiplikation von (4. 1) mit $W(1^{\frac{e}{n}} \cdot K'')$ und Integration ergibt, die Bedingungen

$$(4. 9) \quad \int M \cdot W_u(1^{\frac{e}{n}} \cdot K'') dt = 0$$

für $u = 1 \dots i(1^{\frac{c}{n}} \cdot K'')$ im Fall $i(1^{\frac{c}{n}} \cdot K'') > 0$ zur Existenz von $W(1^{\frac{c}{n}} \cdot K)$ notwendig und, wie ich nachfolgend zeige, auch hinreichend. M braucht

demnach zur Existenz von $W(1^{\frac{c}{n}} \cdot K)$ nicht (4. 8) für $v \neq c$ zu genügen.

Also läßt sich $W(1^{\frac{c}{n}} \cdot K)$ nicht in der Form (4. 7) darstellen, wenn zwar (4. 9),

aber nicht (4. 8) für $v \neq c$ erfüllt und $i(K^{(n)'}) > i(1^{\frac{c}{n}} \cdot K'')$ ist. (4. 7) kann daher nicht die allgemeine Lösung von (4. 1) sein. Jedoch ist nach (4. 1) und

Satz II $W(1^{\frac{c}{n}} \cdot K)$ in $W(K^{(n)})$ für $M_n = T_{n,c}(M)$ enthalten. Demnach ersetze man in (4. 8) M durch $M_n = T_{n,c}(M)$ und beachte

$$(4. 10) \quad \int K^{(e)} \cdot W_u(1^{\frac{v}{n}} \cdot K'') dt = 1^{-\frac{v}{n}} \cdot \int K^{(e-1)} \cdot W_u(1^{\frac{v}{n}} \cdot K'') dt \\ = 1^{-\frac{v}{n}} \cdot W_u(1^{\frac{v}{n}} \cdot K'')$$

Folglich ist

$$\int T_{n,c}(M) \cdot W_u(1^{\frac{v}{n}} \cdot K'') dt = \Sigma_e 1^{\frac{e-v}{n}} \cdot \int M \cdot W_u(1^{\frac{v}{n}} \cdot K'') dt$$

für $e = 0 \dots (n-1)$, somit gleich Null für $c \neq v$ und gleich

$n \cdot \int M \cdot W_u(1^{\frac{v}{n}} \cdot K'') dt$ für $c = v$. Also sind die Bedingungen (4. 8) mit $M_n = T_{n,c}(M)$ anstatt M für $v \neq c$ identisch erfüllt und bleiben für $v = c$

bestehen. Man hat demnach in (4. 9) die $i(1^{\frac{c}{n}} \cdot K'')$ notwendigen und hinreichenden Bedingungen zur Existenz von $W(K^{(n)})$ für $M_n = T_{n,c}(M)$, die

für $i(1^{\frac{c}{n}} \cdot K'') = 0$ wegfallen. Nach Satz II ist

$$(4. 11) \quad A_c = W(K^{(n)}) - 1^{\frac{c}{n}} \cdot \int K \cdot W(K^{(n)})' dt' - M \text{ für } M_n = T_{n,c}(M)$$

eine Lösung von

$$(4. 12) \quad T_{n,c}(W) = 0.$$

Also ist A_c durch die Eigenlösungen von (4. 12) darstellbar. Zu diesen ge-

hören (nach Satz II) die $W_u(1^{\frac{v}{n}} \cdot K)$ für $v \neq c$. Ferner ist jede Eigenlösung von (4. 12) auch eine von (1. 2). Daher kommen (nach Satz III) nur noch die

$W_u(1^{\frac{c}{n}} \cdot K)$ in Betracht, für die aber nach (4. 10)

$$T_{n,c}(W_u(1^{\frac{c}{n}} \cdot K)) = n \cdot W_u(1^{\frac{c}{n}} \cdot K)$$

gilt. Also hat (4. 12) den Defekt $i(K^{(n)}) - i(1^{\frac{c}{n}} \cdot K)$ und die Eigenlösungen

$W_u(1^{\frac{v}{n}} \cdot K)$ für $u = 1 \dots i(1^{\frac{v}{n}} \cdot K)$ und $v = 0 \dots (n-1)$ außer $v = c$.

Folglich gibt es gewisse, durch die Wahl von $W(K^{(n)})$ bestimmte Konstanten p_{cuv} , so daß

$$(4.13) \quad A_e = \sum_v \sum_u p_{cuv} \cdot W_u(1^{\frac{v}{n}} \cdot K)$$

für $u = 1 \dots i(1^{\frac{v}{n}} \cdot K)$ und $v = 0 \dots (n-1)$ außer $v = c$ ist. Führt man hierin

$$(4.13') \quad (1 - 1^{\frac{e-v}{n}}) \cdot W_u(1^{\frac{v}{n}} \cdot K) = W_u(1^{\frac{v}{n}} \cdot K) - 1^{\frac{e}{n}} \cdot \int K \cdot W_u(1^{\frac{v}{n}} \cdot K)' dt'$$

ein, so ergibt sich aus (4.11, 13)

$$\left(W(K^{(n)}) - \sum_v \sum_u \frac{p_{cuv}}{1 - 1^{\frac{e-v}{n}}} \cdot W_u(1^{\frac{v}{n}} \cdot K) \right) - 1^{\frac{e}{n}} \cdot \int K \cdot (\dots)' dt' - M = 0,$$

somit

$$(4.14) \quad W(1^{\frac{e}{n}} \cdot K) = W(K^{(n)}) - \sum_v \sum_u \frac{p_{cuv}}{1 - 1^{\frac{e-v}{n}}} \cdot W_u(1^{\frac{v}{n}} \cdot K)$$

für $u = 1 \dots i(1^{\frac{v}{n}} \cdot K)$, $v = 0 \dots (n-1)$ außer $v = c$ und für $M_n = T_{ne}(M)$, soweit $i(1^{\frac{v}{n}} \cdot K) > 0$ besteht. Dadurch ist $W(1^{\frac{e}{n}} \cdot K)$ aus dem allgemeinen $W(K^{(n)})$ für $M_n = T_{ne}(M)$ ausgesondert. In diesem $W(K^{(n)})$ kann nur noch ein willkürlicher linearer Ausdruck in $W_u(1^{\frac{v}{n}} \cdot K)$ vorkommen. Ich bringe jetzt (4.14) auf die Form (2.3), um dadurch das behauptete (4.6) zu erhalten. Nach (2.3) ist

$$(4.15) \quad W(K^{(n)}) = M - \int \left(- \sum_e 1^{\frac{e}{n}} \cdot K^{(e)} + L(K^{(n)}) + \sum_e 1^{\frac{e}{n}} \cdot \int L(K^{(n)}; t, z) \cdot K^{(e)}(z, t') dz \right) M' dt'$$

für $e = 1 \dots (n-1)$ und $M_n = T_{ne}(M)$, da die vorkommenden Integrationsfolgen vertauscht werden dürfen. Damit geht (4.11) in

$$A_e = \int B_e \cdot M' dt'$$

über; hierin ist

$$\begin{aligned} B_e = & -K^{(n)} - L(K^{(n)}) - \\ & - \sum_e 1^{\frac{e}{n}} \cdot \int L(K^{(n)}; t, z) \cdot K^{(e)}(z, t') dz + \\ & + 1^{\frac{e}{n}} \cdot \int K(t, z) \cdot L(K^{(n)}; z, t') dz + \\ & + \sum_e 1^{\frac{e+1}{n}} \cdot \iint K(t, z) \cdot L(K^{(n)}; z, w) \cdot K^{(e)}(w, t') dw dz. \end{aligned}$$

B_e ist unabhängig von M und hat die gleichen allgemeinen Eigenschaften wie $K^{(n)}$, da das auch für $L(K^{(n)})$ gilt. Somit folgt aus (4. 13)

$$\int B_e \cdot M' dt' = \Sigma_e \Sigma_n p_{euv} \cdot W_u \left(1^{\frac{v}{n}} \cdot K\right).$$

Nimmt man von vornherein die $W_u \left(1^{\frac{v}{n}} \cdot K\right)$ als unitäres System an (Nr. 4a), so ist

$$p_{euv} = \iint \tilde{W}_u \left(1^{\frac{v}{n}} \cdot K\right) \cdot B_e \cdot M' dt' dt.$$

Ich setze

$$H'_{euv} = - \frac{1}{1 - 1^{\frac{e-v}{n}}} \cdot \int \tilde{W}_u \left(1^{\frac{v}{n}} \cdot K\right) \cdot B_e dt$$

und erhalte

$$(4. 16) \quad - \frac{p_{euv}}{1 - 1^{\frac{e-v}{n}}} = \int H'_{euv} \cdot M' dt'.$$

Führt man (4. 15, 16) in (4. 14) ein, so erhält man wegen (2. 3) die behauptete Formel (4. 6). Verzichtet man darauf, daß die $W_u \left(1^{\frac{v}{n}} \cdot K\right)$ ein unitäres System sind, so läßt sich in ähnlicher Weise auf die Existenz der H_{euv} in (4. 6) schließen.

g. Aus (4. 6) leite ich zwei Darstellungen von $L \left(1^{\frac{e}{n}} \cdot K\right)$ ohne H_{euv} ab. Dabei sei g wie in (4. 4) bestimmt. Dann gilt

$$(4. 17) \quad L(K) = - \Sigma_e K^{(e)} + L(K^{(g)}) + \\ + \Sigma_e \int L(K^{(g)}; t, z) \cdot K^{(e)}(z, t') dz$$

für $e = 1 \dots (n-1)$ und $i(K) \geq 0$.

Versieht man von vornherein K mit dem komplexen Faktor $1^{\frac{r}{s}}$, so gewinnt man hieraus eine entsprechende Formel für $L \left(1^{\frac{r}{s}} \cdot K\right)$ an Stelle von (4. 17), wenn $i \left(1^{\frac{r}{s}} \cdot K\right) \geq 0$ ist. Andererseits gilt

$$(4. 18) \quad L \left(1^{\frac{e}{n}} \cdot K\right) = - \Sigma_e 1^{\frac{e}{n}} \cdot K^{(e)} + L(K^{(n)}) + \\ + \Sigma_e 1^{\frac{e}{n}} \cdot \int L(K^{(n)}; t, z) \cdot K^{(e)}(z, t') dz$$

für $i(K^{(n)}) = 0$ und damit gleichbedeutend für $i(K) = 0$ und $n = g$.

Hievon ist der Sonderfall

$$L(K) = - \Sigma_e K^{(e)} + L(K^{(n)}) + \Sigma_e \int L(K^{(n)}; t, z) \cdot K^{(e)}(z, t') dz$$

für $e = 1 \dots (n-1)$ und $i(K^{(n)}) = 0$ bekannt. Das kann aus der bei Fredholm anschließend (27) stehenden Formel gefolgert werden. Es ist zu beachten,

daß bei Fredholm K und L das umgekehrte Vorzeichen wie hier haben. Also ergibt sich nach Fredholm, wenn ich $K = -K$ und $L(K) = -L(K)$ setze,

$$L(K) = L(-(-1)^n \cdot K^{(n)}) + \sum_e (-1)^e \cdot K^{(e)} + \\ + \int L(-(-1)^n \cdot K^{(n)}; t, z) \cdot \sum_e (-1)^e \cdot K^{(e)}(z, t') dz \\ \text{für } e = 1 \dots (n-1).$$

[Vgl. ferner Fredh., Formel nach (26) und vor (25) daselbst.] Diese Vorzeichenwahl hat auch Hellinger-Toeplitz in der Auflösungstheorie entgegen der in der Eigenwerttheorie, die mit meiner übereinstimmt, getroffen. Dann ist Formel (3) auf S. 1384 in Hell.-Toepl. durch obige Formel für $L(K)$ zu berichtigen: K_n ist nicht der lösende Kern von $K^{(n)}$, sondern von $-(-1)^n \cdot K^{(n)}$. Weiter ist beim Vergleich der Formel in Hell.-Toepl. mit obigem $L(K)$ noch die spätere Formel (7. 10) zu berücksichtigen.

h. Ich beweise

$$(4.19) \quad L(K^{(n)}) = L(K) + L(1^{\frac{1}{n}} \cdot K) + \\ + \dots + L(1^{\frac{n-1}{n}} \cdot K) - \int L(K; t, s) \cdot L(1^{\frac{1}{n}} \cdot K; s, t') ds - \\ - \int L(K; t, s) \cdot L(1^{\frac{2}{n}} \cdot K; s, t') ds - \\ - \dots - \int L(1^{\frac{n-2}{n}} \cdot K; t, s) \cdot L(1^{\frac{n-1}{n}} \cdot K; s, t') ds + \\ + \iint L(K; t, s) \cdot L(1^{\frac{1}{n}} \cdot K; s, w) \cdot L(1^{\frac{2}{n}} \cdot K; w, t') ds dw + \\ + \iint L(K; t, s) \cdot L(1^{\frac{1}{n}} \cdot K; s, w) \cdot L(1^{\frac{3}{n}} \cdot K; w, t') ds dw \\ + \dots + \iint L(1^{\frac{n-3}{n}} \cdot K; t, s) \cdot L(1^{\frac{n-2}{n}} \cdot K; s, w) \cdot L(1^{\frac{n-1}{n}} \cdot K; w, t') ds dw \\ \text{usw. bis} \\ + (-1)^{n-1} \cdot \iint \dots \int L(K; t, s) \times \\ \times L(1^{\frac{1}{n}} \cdot K; s, w) \dots L(1^{\frac{n-1}{n}} \cdot K; z, t') ds dw \dots dz$$

für $i(K^{(n)}) = 0$ und damit gleichbedeutend für $i(K) = 0$ und $n = g$ wie in (4. 4), wenn $n \geq 2$ ist. Dabei sind alle $L(1^{\frac{n}{n}} \cdot K)$ nach Nr. 4f und g vorhanden.

Man setze nacheinander

$$M_{v+1} = M_v - \frac{1}{n} \cdot \int K \cdot M'_v dt'$$

für $v = 0 \dots (n-1)$ und $M_0 = M$. Dann gilt nach Satz I

$$M_n = M - \int K^{(n)} \cdot M' dt',$$

also nach (2.3)

$$(4.20) \quad M = M_n - \int L(K^{(n)}) \cdot M'_n dt'.$$

Geht man in der rechten Seite von (4.19) bis a anstatt n , $1 \leq a \leq n$, und nennt den so entstehenden Ausdruck C_a , so gilt

$$(4.21) \quad M = M_a - \int C_a \cdot M'_a dt'$$

vorerst für $a = 1$, wenn man $C_1 = L(K)$ setzt. Angenommen, (4.21) gelte außerdem für irgendein a , das $2 \leq a \leq n-1$ genügt, falls $n > 2$ ist. Setzt man dann in (4.21) für $1 \leq a \leq n-1$ die gemäß (2.3) bestehende Gleichung

$$M_a = M_{a+1} - \int L(1^{\frac{a}{n}} \cdot K) \cdot M'_{a+1} dt'$$

ein, so erhält man (4.21) für $(a+1)$ anstatt a ; denn es ist

$$C_a + L(1^{\frac{a}{n}} \cdot K) - \int C_a(t, s) \cdot L(1^{\frac{a}{n}} \cdot K; s, t') ds = C_{a+1}.$$

Also besteht (4.21) auch für $a = n$, somit nach (4.20) $C_n = L(K^{(n)})$, das heißt (4.19). Andererseits folgt aus (4.18) die Formel

$$(4.22) \quad n \cdot L(K^{(n)}) = \sum_c L(1^{\frac{c}{n}} \cdot K)$$

für $c = 0 \dots (n-1)$; wenn $i(K^{(n)}) = 0$ und damit gleichbedeutend $i(K) = 0$ und $n = g$ wie in (4.4) ist. Demnach ist in Hell-Toepl. die Formel (4) auf S. 1384 dadurch zu berichtigen, daß, wie schon in Nr. 4g festgestellt, K_n der lösende Kern von $-(-1)^n \cdot K^{(n)}$ sein muß und daß ferner alle ε -Potenzen durch 1 zu ersetzen sind. Durch Vergleich der beiden Darstellungen von $L(K^{(n)})$ erhält man eine Integralbeziehung zwischen den $L(1^{\frac{c}{n}} \cdot K)$, beispielsweise

$$L(K) + L(-K) = 2 \cdot \int L(K; t, s) \cdot L(-K; s, t') ds$$

für $n = 2$, wenn $i(K^{(2)}) = 0$ ist.

5. Beweis der Hauptsätze der Auflösungstheorie mittels Iteration der ganzen Integralgleichung.

a. Der Inhalt der Nr. 4 besagt vom Standpunkt der Auflösungstheorie aus gesehen, daß die *Fredholmsche Integralgleichung im engeren Sinne un-eigentlich singular* ist. Für sie gelten also die Hauptsätze. Bevor ich darauf näher eingehe, bemerke ich, daß K einen komplexen Zahlenfaktor enthalten darf und daß die Voraussetzung III der Nr. 1 betreffs Unstetigkeitsstellen

bisher nicht benutzt wurde. Ich könnte demnach die Definition der Fredholmschen Gleichung noch etwas weiter fassen. Die Aufstellung der Voraussetzungen I bis IV geschieht im Hinblick auf die (in Nr. 6 und 7 entwickelte) Theorie des allgemeinsten lösenden Kernes.

Im einzelnen stelle ich fest: Die Endlichkeit von jedem $i(1^n \cdot K)$, das heißt der erste Defektsatz, folgt aus (4. 3). Der zweite Defektsatz ist (4. 5). Der Existenzsatz für den lösenden Kern ergibt sich aus der Darstellung von $L(1^n \cdot K)$ durch K und $L(K^{(n)})$ in (4. 6) und ebenso aus der in (4. 17) für $1^n \cdot K$ anstatt K ; beidem zufolge ist $L(1^n \cdot K)$ vorhanden und hat die gleichen allgemeinen Eigenschaften wie K . Der Lösungssatz gilt, da mit Rücksicht auf das vor (4. 11) Festgestellte (4. 14) dann und nur dann besteht, wenn M den Bedingungen (4. 9) im Fall $i(1^n \cdot K'') > 0$ genügt. Für $i(1^n \cdot K'') = 0$ fallen diese Bedingungen weg. Man kann erreichen, daß in der Lösung (4. 14) die $p_{\epsilon u v}$ nicht auftreten, wenn man daselbst wie in der Formel (4. 17) und ihrem Nachsatz zuerst $c = 0$ setzt, weiter $n = g$ wie in (4. 4) wählt und schließlich K durch $1^n \cdot \tilde{K}$ ersetzt.

b. Die als gemeinsamer Sonderfall der Fredholmschen und Levischen Gleichungen angekündigte Integralgleichung (1. 1) mit $K = N : R^p$ ist eine Fredholmsche Gleichung, weil es (nach Nr. 3c) ein m gibt, für das $K^{(n)}$ im Fall $n \geq m$ stetig ist. Denn dann gelten nach dem ersten Satz der Nr. 2c die Hauptsätze für (1. 2) und ist daher die Voraussetzung IV der Nr. 1 erfüllt. Folglich ist auch der Sonderfall mit $K = N : R^p$ im engeren Sinn uneigentlich singular. Die Frage, ob es eine Fredholmsche Gleichung $\{ \}^\dagger$, deren K nicht in der Form $N : R^p$ oder nicht durch eine endliche Anzahl solcher Ausdrücke darstellbar ist, ist noch nicht geklärt. Allgemeiner könnte die Singularität von K so beschaffen sein, daß sie sich nur schwer und unübersichtlich oder überhaupt nicht auf Ausdrücke der Art $N : R^p$ zurückführen läßt. Damit hängt zusammen, daß die allgemeine Fredholmsche Gleichung und der Sonderfall mit $K = N : R^p$ beide im engeren Sinn uneigentlich singular sind und daß auch sonst im Hinblick auf die Auflösungstheorie kein wesentlicher Unterschied bekannt ist.

e. Ich beweise noch eine Fredholmsche Determinantengleichung, wie ich in Nr. 2d IV angekündigt habe. Die Zusammensetzung der beiden Integraltransformationen $W - 1^n \cdot \int K \cdot W' dt'$ und $W - \int K^{(n)} \cdot W' dt'$ ist unabhängig von der Reihenfolge. Demnach gibt es gewisse Konstanten f_{uv} und f'_{uv} , für die (2. 8, 9) im Fall $i(K^{(n)}) > 0$ bestehen. Multipliziert man (2. 8) mit

$W_v(K''^{(n)})$ und (2.9) mit $W_u(K^{(n)})$, integriert über t und subtrahiert voneinander, so erhält man

$$(5.1) \quad \sum_r f_{ur} \cdot \int W_r(K^{(n)}) \cdot W_v(K''^{(n)}) dt = \sum_s f'_{uc} \cdot \int W_u(K^{(n)}) \cdot W_s(K''^{(n)}) dt$$

Also ist

$$\begin{aligned} & (\text{Determinante der } f_{uv}) \times (\text{Det. der } \int W_u(K^{(n)}) \cdot W_v(K''^{(n)}) dt) \\ &= (\text{Det. der } f'_{uv}) \times (\text{Det. der } \int W_u(K^{(n)}) \cdot W_v(K''^{(n)}) dt), \end{aligned}$$

weil $i(K^{(n)}) = i(K''^{(n)})$ unter der Voraussetzung $n \geq m$ ist. Folglich gilt (2. falls die Determinante der $\int W_u(K^{(n)}) \cdot W_v(K''^{(n)}) dt$ ungleich Null ist. Darüber hinaus beweise ich die Möglichkeit einer geeigneten Numerierung der f_{uv} und f'_{uv} bei der sogar

$$(5.2) \quad f_{uv} = f'_{uv}$$

und damit insbesondere (2.7) in allen Fällen gilt. Es sei $i(1^n \cdot K) > 0$ für $a = a_1 \dots a_z$ und $i(1^n \cdot K) = 0$ für alle übrigen Werte von a der Reihe $0 \dots (n-1)$. Demnach ist $z \leq n$. Nach Satz III der Nr. 4 setze ich

$$W_u(K^{(n)}) = W_u(1^{\frac{a_1}{n}} \cdot K) \text{ für } u = 1 \dots i(1^{\frac{a_1}{n}} \cdot K),$$

$$W_u(K^{(n)}) = W_u(1^{\frac{a_2}{n}} \cdot K) \text{ für } u = (i(1^{\frac{a_1}{n}} \cdot K) + 1) \dots (i(1^{\frac{a_1}{n}} \cdot K) + i(1^{\frac{a_2}{n}} \cdot K))$$

usw. bis

$$W_u(K^{(n)}) = W_u(1^{\frac{a_z}{n}} \cdot K) \text{ für } u = (\sum_s i(1^{\frac{a_s}{n}} \cdot K) + 1) \dots (\sum_s i(1^{\frac{a_s}{n}} \cdot K) + i(1^{\frac{a_z}{n}} \cdot K)),$$

$$s = 1 \dots (z-1).$$

Dabei ist

$$\sum_s i(1^{\frac{a_s}{n}} \cdot K) + i(1^{\frac{a_z}{n}} \cdot K) = i(K^{(n)});$$

ferner habe ich die unteren Indizes der $W_u(1^{\frac{a}{n}} \cdot K)$ umbezeichnet. Setze ich diese Werte in die Identität (4.13') für $v = a_1 \dots a_z$ ein und vergleiche mit der Definitionsgleichung (2.8) für f_{uv} , so erhalte ich

$$(5.3) \quad f_{uv} = 0 \text{ für } u \neq v,$$

$$f_{uu} = 1 - 1^{\frac{c-a_1}{n}} \text{ für } u = 1 \dots i(1^{\frac{a_1}{n}} \cdot K),$$

$$f_{uu} = 1 - 1^{\frac{c-a_r}{n}} \text{ für } u = (\sum_s i(1^{\frac{a_s}{n}} \cdot K) + 1)$$

$$\text{bis } (\sum_s i(1^{\frac{a_s}{n}} \cdot K) + i(1^{\frac{a_r}{n}} \cdot K)),$$

$$s = 1 \dots (r-1) \text{ und } r = 2 \dots z.$$

Entsprechendes gilt im transponierten Fall, bei dem a'_v und z'' an die Stelle von a_u und z treten mögen. Man kann dann nach (4.5) stets $a_v = a'_v$ und $z = z''$ setzen, woraus (5.2) und damit (2.7) folgt.

Ich bemerke noch, daß die Determinante der f_{uv} gleich dem Produkt aller derjenigen $(1 - 1^{\frac{c-\alpha}{n}})^i$ ist, für die $i = i(1^{\frac{\alpha}{n}} \cdot K) > 0$ ist. Sie ist also genau dann ungleich Null, wenn $i(1^{\frac{\alpha}{n}} \cdot K) = 0$ ist. Weiter folgt aus (5.1, 2, 3)

$$(f_{uu} - f_{vv}) \cdot \int W_u(K^{(n)}) \cdot W_v(K''^{(n)}) dt = 0.$$

Hiermit sind beide Fredholmsche Annahmen betreffs der Gültigkeit von (2.7) verträglich, nämlich daß die Determinante der $\int W_u(K^{(n)}) \cdot W_v(K''^{(n)}) dt$ ungleich Null oder gleich Null ist.

6. Eigenschaften des allgemeinsten lösenden Kernes.

Zur Angabe des allgemeinsten lösenden Kernes L und seiner Eigenschaften, ferner der kennzeichnenden Sonderfälle der Theorie des lösenden Kernes sollen in den nachfolgenden Nrn. 6 und 7 für die Ausgangsintegralgleichung (1.1) die Voraussetzungen I bis III der Nr. 1 erfüllt sein und die Hauptsätze gelten. Wiederholt schreibe ich von irgendeiner Funktion $C = C(t, t')$ bzw. $A = A(t)$, sie sei von der Art wie K bzw. M und meine damit, daß sie die Voraussetzungen I bis III der Nr. 1 erfüllt, wenn man daselbst K durch C bzw. M durch A ersetzt.

Zwei lösende Kerne, die sich nur an ihren Unstetigkeitsstellen unterscheiden, stellen (nach Voraussetzung III der Nr. 1) das gleiche W dar. Ich bezeichne sie daher als übereinstimmend.

a. Ich zeige zuerst, daß L_{00} sich als den (im oben angegebenen Sinn) einzigen lösenden Kern bestimmen läßt, für den

$$(6.1) \quad W_u' - \int L_{00} \cdot W_u dt = 0, \quad u = 1 \dots i,$$

und

$$(6.2) \quad W_v'' - \int L_{00} \cdot W_v'' dt' = 0, \quad v = 1 \dots i'',$$

in b besteht.

Denn ist L irgendein lösender Kern und bestimmt man

$$\begin{aligned} A \equiv & L + \sum_u W_u \cdot (W_u' - \int W_u \cdot L dt) + \\ & + \sum_v W_v'' \cdot (W_v'' - \int W_v'' \cdot L dt') - \\ & - \sum_u \sum_v W_u \cdot W_v'' \cdot (\int W_u \cdot W_v'' dt - \int \int W_u \cdot W_v'' \cdot L dt dt') \end{aligned}$$

für $u = 1 \dots i$ und $v = 1 \dots i''$,

so ist A wie L unabhängig von M und von der Art wie K . Ferner ist, wenn man (2. 1) voraussetzt,

$$M - \int A \cdot M' dt' = M - \int L \cdot M' dt' + \sum_u W_u \cdot h_u;$$

hierin sind die h_u gewisse Konstanten. Also ist $M - \int A \cdot M' dt'$ eine gewisse Lösung von (1. 1), da (2. 4) auch für L anstatt L_{00} gilt. Folglich ist A ein lösender Kern, der (6. 1, 2) für A anstatt L_{00} erfüllt. Angenommen, C sei irgendein lösender Kern, der (6. 1, 2) für C anstatt L_{00} befriedigt. Dann besteht (2. 4) ebenso für A als auch für C anstatt L_{00} , wobei die beiden Systeme k_u verschieden sein können. Hieraus folgt $\int (A - C) \cdot M' dt' = \sum_u k_u \cdot W_u$, somit $k_u = 0$ für $u = 1 \dots i$, wenn man

$$\int A \cdot W_u dt = \int C \cdot W_u dt (= W_u')$$

berücksichtigt. Wegen

$$\int A \cdot W_v'' dt' = \int C \cdot W_v'' dt' (= W_v'')$$

wird (2. 1) durch $(A - C)$ anstatt M' erfüllt. Aber ich kann im allgemeinen nicht durchweg $M' = A - C$ setzen; denn es ist unbekannt, ob $(A - C)^2$ überall integrierbar ist. Nach Voraussetzung III der Nr. 1 gibt es einen Teilbereich B von b , der alle Unstetigkeitsstellen von $(A - C)$ enthält. Dann darf ich

$$M' = (A - C) - \sum_v W_v'' \cdot \int_{b-B} (A - C) \cdot W_v'' dt' \quad \text{für } t' \text{ in } (b - B)$$

und

$$M' = - \sum_v W_v'' \cdot \int_{b-B} (A - C) \cdot W_v'' dt' \quad \text{für } t' \text{ in } B$$

setzen. Dieses M' befriedigt (2. 1) und ergibt in $\int (A - C) \cdot M' dt' = 0$ eingesetzt $\int_{b-B} (A - C)^2 dt' = 0$, also $A - C = 0$ für t in b und t' in $(b - B)$.

Jetzt kann ich (nach Voraussetzung III der Nr. 1) B so klein machen, daß es nur noch die Unstetigkeitsstellen von A und C enthält. Da es auf die Werte an den Unstetigkeitsstellen nicht ankommen soll, ist demnach $L_{00} = A = C$ der einzige lösende Kern, für den (6. 1, 2) besteht. Das ist allein kennzeichnend für L_{00} . Alle anderen Eigenschaften sind in denen für die nachfolgend erklärten lösenden Kerne, deren *Grundbestandteil* L_{00} ist, enthalten.

b. Der *allgemeinste lösende Kern* soll jede Funktion L enthalten, für die (2. 3) eine Lösung von (1. 1) ist; dabei ist L unabhängig von M und von der Art wie K . Es seien $A_u = A_u(t)$ für $u = 1 \dots i$ und $B_v = B_v(t)$ für $v = 1 \dots i''$ willkürlich vorgegebene Funktionen der Art wie M . Ich beweise

$$(6. 3) \quad L = L_{00} + \sum_u W_u \cdot A_u' + \sum_v B_v \cdot W_v''$$

für $u = 1 \dots i$ und $v = 1 \dots i''$.

Alle L mit gleichen A_u ergeben das gleiche W , wie aus (2. 3, 1) folgt.

Denn denkt man sich zuerst L durch (6.3) definiert, so ist L wie L_{00} von der Art wie K , ferner $M - \int L \cdot M' dt'$ wegen (2.1, 4) eine Lösung von (1.1), also L nach (2.3) ein lösender Kern. Ist andererseits L irgendein lösender Kern, so ist nach (2.3, 4) $\int (L_{00} - L) \cdot M' dt' = \sum_u k_u \cdot W_u$, somit $\iint (L_{00} - L) \cdot M' \cdot W_u dt' dt = k_u$. Setzt man

$$A_u' = - \int (L_{00} - L) \cdot W_u dt,$$

so ist $\int (L_{00} + \sum_u W_u \cdot A_u' - L) \cdot M' dt' = 0$; die A_u sind unabhängig von M . Damit ist wegen (2.1) auch $\int (L_{00} + \sum_u W_u \cdot A_u' + \sum_v B_v \cdot W_v'' - L) \times M' dt' = 0$; hierin kann man über die B_v (im Rahmen der allgemeinen Voraussetzungen) noch frei verfügen. Die B_v bestimme ich so, daß

$$\int (L_{00} + \sum_u W_u \cdot A_u' + \sum_v B_v \cdot W_v'' - L) \cdot W_v'' dt' = 0$$

ist, demnach

$$B_v = - \int (L_{00} + \sum_u W_u \cdot A_u' - L) \cdot W_v'' dt';$$

auch die B_v sind unabhängig von M . Man wähle (ähnlich wie in Nr. 6a)

$$M' = (L_{00} + \sum_u W_u \cdot A_u' + \sum_v B_v \cdot W_v'' - L) - \sum_v W_v'' \cdot \int_{b-B} (\dots) \cdot W_v'' dt' \text{ für } t' \text{ in } (b-B)$$

und

$$M' = - \sum_v W_v'' \cdot \int_{b-B} (\dots) \cdot W_v'' dt' \text{ für } t' \text{ in } B;$$

hierin enthält B alle Unstetigkeitsstellen von (\dots) . Dann ergibt sich $(\dots) = 0$ für t in b und t' in $(b-B)$, somit (6.3) für die vorstehend bestimmten A_u und B_v ; denn es soll auf die Werte an den Unstetigkeitsstellen nicht ankommen. Daher ist $(L_{00} + \sum_u W_u \cdot A_u' + \sum_v B_v \cdot W_v'')$ ein lösender Kern und umgekehrt jeder lösende Kern in dieser Summe enthalten. Das gerade drückt (6.3) aus.

c. Es seien $F_v'' = F_v''(t)$ für $v = 1 \dots i''$ Funktionen der Art wie M , die

$$(6.4) \quad \int (F_v'' - W_v'') \cdot W_u'' dt = 0$$

für u und $v = 1 \dots i''$ genügen und im übrigen beliebig vorgegeben sind. Ferner sei $Z = Z(t, t')$ eine von M unabhängige Funktion der Art wie K , die der Integralgleichung

$$(6.5) \quad Z - \int K(t, s) \cdot Z(s, t') ds = -K + \sum_v F_v'' \cdot W_v''$$

genügen soll, wenn man t' von den Unstetigkeitsstellen ausschließt.

Ich beweise den Satz: (6.5) wird durch $Z = L$ gelöst, wenn man die F_v'' aus vorgegebenen B_v mittels

$$(6.6) \quad F_v'' = W_v'' + B_v - \int K \cdot B_v' dt'$$

bestimmt, wobei L durch (6.3) gegeben ist; andererseits gibt es stets ein L der Form (6.3), das (6.5) mit vorgegebenen F_v'' , die (6.4) erfüllen, löst; umgekehrt ist jedes zulässige Z in L enthalten. Die A_u bleiben dabei stets willkürlich. Die F_v'' können für $i'' > 0$ nie identisch verschwinden. Damit ist (6.5) eine *erste kennzeichnende Integralgleichung für den lösenden Kern*. Entsprechendes gilt im transponierten Fall.

Setzt man zum Beweis $W = M - \int L_{00} \cdot M' dt'$ in (1.1) ein, so erhält man

$$\int (L_{00} - \int K(t, s) \cdot L_{00}(s, t') ds + K) \cdot M' dt' = 0,$$

somit wegen (2.1)

$$\int (L_{00} - \int K(t, s) \cdot L_{00}(s, t') ds + K - \sum_v F_v'' \cdot W_v''') \cdot M' dt' = 0$$

für beliebige F_v'' .

Ich wähle die F_v'' so, daß $\int (\dots) \cdot W_v''' dt' = 0$ gilt, also wegen (6.2) $F_v'' = W_v''$. Man findet (nach einem in Nr. 6a und b angewandten Verfahren)

$$(6.7) \quad L_{00} - \int K(t, s) \cdot L_{00}(s, t') ds = -K + \sum_v W_v'' \cdot W_v''',$$

wenn man t' von den Unstetigkeitsstellen ausschließt. Setzt man das durch (6.3) bestimmte L in (6.5) ein, so erhält man (6.6) wegen (6.7), weil die W_v'' untereinander linear unabhängig sind. Also wird (6.5) durch $Z = L$ gelöst, wenn man die F_v'' durch (6.6) bestimmt. Diese F_v'' genügen (6.4) auf Grund der Definition der W_v'' . (6.4) sind notwendige Bedingungen für F_v'' zur Existenz von Z ; das folgt aus (6.5) nach Multiplikation mit W_v'' und Integration über t , da die W_v'' untereinander linear unabhängig sind. (6.4) sind ferner hinreichende Bedingungen; denn bestimmt man die B_v als sodann existierende Lösungen von (6.6) aus vorgegebenen F_v'' und bildet L gemäß (6.3) mit diesen B_v auf Grund der Existenz von L_{00} , so wird (6.5) von $Z = L$ gelöst. Ist Z irgendeine zulässige Lösung von (6.5) und multipliziert man (6.5) mit M' und integriert über t' , dann ist wegen (2.1) $M - \int Z \cdot M' dt'$ eine Lösung von (1.1); folglich ist Z nach (2.3) ein lösender Kern und daher in (6.3) enthalten.

d. Nach dem auch im transponierten Fall gültigen Lösungssatz wird

$$(6.8) \quad W'' - \int K'' \cdot W''' dt' = M''$$

durch ein W'' gelöst, wenn M'' die dazu notwendigen und hinreichenden Bedingungen

$$(6.9) \quad \int M'' \cdot W_u dt = 0 \quad \text{für } u = 1 \dots i$$

erfüllt und im übrigen M'' von gleicher Art wie M ist. Multipliziert man (6. 7) mit W'' , integriert über t und beachtet (6. 8), so findet man

$$W'' = M'' - \int L''_{00} \cdot M''' dt' + \sum_v W''_v \cdot \int W''_v \cdot W'' dt.$$

Man setze in $\int W''_v \cdot W'' dt$ für W'' die (2. 4) entsprechende Darstellung ein. Dann erhält man für $\int W''_v \cdot W'' dt$ gewisse Konstanten, die durch W'' bestimmt sind. Entsprechend wie für (1. 1) ist mit W'' auch $W'' - \sum_v W''_v \cdot \int W''_v \times W'' dt$ eine Lösung von (6. 8). Folglich ist L''_{00} ein lösender Kern zu K'' . Ferner genügt L''_{00} den zu (6. 1, 2) transponierten Integralbedingungen, weil (6. 1, 2) bei Vertauschung von (t, t') mit (t', t) in die transponierten Formen übergehen. Demnach ist L''_{00} der L_{00} entsprechende, eindeutige spezielle lösende Kern zu K'' , das heißt

$$(6. 10) \quad L_{00}(K'') = L_{00}(K)'.$$

Somit ist das nach (6. 3) zu bestimmende L'' der allgemeinste lösende Kern zu K'' , das heißt

$$(6. 11) \quad L(K'') = L(K)'.$$

Folglich ist

$$W'' = M'' - \int L''_{00} \cdot M''' dt' + \sum_v k'_v \cdot W''_v$$

für $v = 1 \dots i''$ mit den willkürlichen Konstanten k'_v die allgemeine Lösung von (6. 8), wenn (6. 9) besteht. Alle L'' mit gleichen B_v stellen wegen (6. 9) das gleiche W'' dar.

e. Der Satz der Nr. 6c gilt auch im transponierten Fall, den man dann auf Grund von (6. 11) als einen zweiten Satz für L ausdrücken kann. Hiervon merke ich allein eine *zweite kennzeichnende Integralgleichung für den lösenden Kern* an:

$$(6. 12) \quad \begin{aligned} L - \int L(t, s) \cdot K(s, t') ds &= -K + \sum_u W_u \cdot F_u', \\ F_u' &= W_u' + A_u' - \int K \cdot A_u dt, \\ \int (F_u - W_u) W_v dt &= 0, \\ u \text{ und } v &= 1 \dots i. \end{aligned}$$

Dabei ist das über die Unstetigkeitsstellen wiederholt Geschriebene zu beachten.

Demnach hat L drei untereinander gleichwertige kennzeichnende Eigenschaften, von denen jede als Definition benutzt werden kann, und aus der dann die beiden anderen sich ergeben:

- I. L ist der allgemeinste Kern der Darstellung (2. 3).
- II. L ist die allgemeine Lösung der Integralgleichung (6. 5).
- III. L ist die allgemeine Lösung der Integralgleichung (6. 12).

7. Sonderfälle der Theorie des lösenden Kernes.

a. Für einen Sonderfall des lösenden Kernes, den ich mit $L_0 = L_0(t, t')$ bezeichne, gebe ich fünf untereinander gleichwertige kennzeichnende Eigenschaften an. Von diesen bevorzuge ich die erste als Definition.

I. Ich definiere

$$(7.1) \quad L_0 = L_{00} + \sum_u \sum_v k_{uv} \cdot W_u \cdot W_v'''$$

für $u = 1 \dots i$ und $v = 1 \dots i''$ mit willkürlich vorgegebenen Konstanten k_{uv} . L_0 umfaßt L_{00} und ist gemäß (6.3) in L enthalten. L_0'' hat für K'' die gleiche Bedeutung wie L_0 für K , das heißt

$$(7.2) \quad L_0(K'') = L_0(K)'.$$

II. L_0 sind diejenigen und nur diejenigen L , welche die Integralbedingungen

$$(7.3) \quad W_u' - \int L_0 \cdot W_u dt = - \sum_v k_{uv} \cdot W_v'''$$

und

$$(7.4) \quad W_v''' - \int L_0 \cdot W_v''' dt' = - \sum_u k_{uv} \cdot W_u$$

erfüllen. Denn einerseits folgen sie aus (6.1, 2). Umgekehrt sei L irgendein lösender Kern, der (7.3, 4) erfüllt. Nach einem in Nr. 6a und b angewandten Verfahren ist dann $\int (L - L_0) \cdot M' dt' = \sum_u k_u \cdot W_u$, $k_u = 0$, $L - L_0 = 0$; denn es soll (nach Voraussetzung III der Nr. 1) auf die Werte an den Unstetigkeitsstellen nicht ankommen.

III. Ist W_0 irgendeine Lösung der Ausgangsintegralgleichung (1.1) bzw. W_0'' eine der dazu transponierten Integralgleichung (6.8), so läßt sich W_0 bzw. W_0'' stets als einzige Lösung bestimmen, die

$$(7.5) \quad \int W_0 \cdot W_u dt = 0 \quad \text{für } u = 1 \dots i$$

bzw.

$$\int W_0'' \cdot W_v''' dt = 0 \quad \text{für } v = 1 \dots i''$$

genügt (Courant, Bd. 1, S. 102). Ich beweise, daß L_0 diejenigen und nur diejenigen L sind, die W_0 und W_0'' bei erfüllten Lösungsbedingungen (2.1) für M und (6.9) für M'' darstellen, das heißt

$$W_0 = M - \int L_0 \cdot M' dt'$$

und

$$W_0''' = M''' - \int L_0 \cdot M'' dt.$$

L_0 ist zum Unterschied von W_0 und W_0'' nicht eindeutig. Erst die Forderung $k_{uv} = 0$ führt auf das eindeutige L_{00} , das also auch W_0 und W_0'' darstellt. Aus

$$W = M - \int L_0 \cdot M' dt' = M - \int L_{00} \cdot M' dt'$$

folgt wegen (6. 1) und (7. 5) $W = W_0$, da W_0 eindeutig ist. Umgekehrt sei L irgendein lösender Kern zu W_0 , also $W_0 = M - \int L \cdot M' dt'$. Daraus folgt wegen der Eindeutigkeit von W_0 $\int (L - L_{00}) \cdot M' dt' = 0$. Also ist nach (6. 3) und (2. 1), da die W_u untereinander linear unabhängig sind, $\int A_u \cdot M dt = 0$, ferner $\int (A_u + \Sigma_v p_{uv} \cdot W_v'') \cdot M dt = 0$. Hierin wähle ich die Konstanten p_{uv} so, daß $\int (A_u + \Sigma_v p_{uv} \cdot W_v'') \cdot W_u'' dt = 0$ ist. M darf ich dann durch $(A_u + \Sigma_v p_{uv} \cdot W_v'')$ ersetzen, woraus $A_u + \Sigma_v p_{uv} \cdot W_v'' = 0$ folgt. Demnach sind die A_u in L linear abhängig von den W_v'' . Stellt L dazu noch W_0'' dar, so folgt aus $W_0'' = M'' - \int L \cdot M'' dt$ auf entsprechende Weise, daß die B_v in L von den W_u linear abhängen. Also geht (6. 3) in (7. 1) über, das heißt, dieses L ist in L_0 enthalten. Demnach ist obige Behauptung richtig.

IV. Ich beweise: L_0 sind die und nur die Lösungen einer ersten kennzeichnenden Integralgleichung, nämlich von

$$(7. 6) \quad Z - \int K(t, s) \cdot Z(s, t') ds = -K + \Sigma_v W_v'' \cdot W_v''',$$

wenn Z die Nebenbedingungen

$$(7. 7) \quad W_u' - \int Z \cdot W_u dt = -\Sigma_v k_{uv} \cdot W_v'''$$

befriedigt; $u = 1 \dots i$ und $v = 1 \dots i''$; dabei beschränke man $Z = Z(t, t')$ auf die von \bar{M} unabhängigen Funktionen der Art wie K . Insbesondere löst L_{00} diese Gleichung für $k_{uv} = 0$.

Sind in (6. 3) die B_v von den W_u linear abhängig, so erfüllt L die Integralgleichung (6. 5) für $F_v'' = W_v''$, folglich $Z = L_0$ die Gleichung (7. 6). Ferner geht (7. 7) für $Z = L_0$ in (7. 3) über. Umgekehrt ist jede Lösung von (7. 6) in (6. 3) enthalten; denn $F_v'' = W_v''$ erfüllt die Bedingungen (6. 4). Dabei ist $B_v - \int K \cdot B_v' dt' = 0$. Also sind die B_v von den W_u linear abhängig. Aus (6. 3, 1) und (7. 7) für $Z = L$ folgt ferner, daß

$$\begin{aligned} W_u' - \int L \cdot W_u dt &= -A_u' - \Sigma_v \int B_v \cdot W_u dt \cdot W_v''' \\ &= -\Sigma_v k_{uv} \cdot W_v''' \end{aligned}$$

ist. Folglich sind die A_u' von den W_v''' linear abhängig. L geht demnach in L_0 über, womit die Behauptung bewiesen ist.

Verzichtet man jedoch auf (7. 7), so erhält man in (7. 6) ein kennzeichnendes Randwertproblem für einen in L enthaltenen und L_0 umfassenden lösenden Kern. Solche Zwischenbildungen von lösenden Kernen, für die ähnliche Aussagen aufstellbar sind, werden in vorliegender Arbeit nicht weiter behandelt.

V. Entsprechendes gilt im transponierten Fall. Dadurch entsteht wegen (7. 2) eine zweite kennzeichnende Integralgleichung für L_0 mit (7. 4) für Z anstatt L_0 als Nebenbedingungen.

b. Im sich selbst transponierten Fall $K = K''$ ist $L = L''$, $L_0 = L'_0$ und $L_{00} = L''_{00}$.

Denn L'' ist der lösende Kern zu K'' , also auch zu K und somit wieder von der allgemeinen Form L . Die beiden L bzw. L_0 kennzeichnenden Integralgleichungen fallen zusammen.

c. Es sei $C_u = C_u(t)$ eine Eigenlösung von L_{00} , das heißt $C_u - \int L_{00} \cdot C_u' dt' = 0$. Multipliziert man (6. 7) mit C_u' und integriert über t' , so findet man $C_u = \sum_v W_v'' \cdot \int W_v'' \cdot C_u' dt'$. Entsprechend erhält man im transponierten Fall wegen (6. 10), daß jede Eigenlösung von L''_{00} von den W_u linear abhängt. Andererseits sind nach (6. 2, 1) alle W_v'' bzw. W_u Eigenlösungen von L_{00} bzw. L''_{00} . Somit hat L_{00} genau die Eigenlösungen W_v'' für $v = 1 \dots i''$ wie K'' und L''_{00} genau die Eigenlösungen W_u für $u = 1 \dots i$ wie K . Dann sind $W_u - \int K \cdot W_u' dt' = 0$ und $W_v'' - \int K \cdot W_v'' dt = 0$ die (6. 2, 1) entsprechenden Integralbedingungen für K . Ferner lassen sich (6. 7) und die entsprechende Integralgleichung für $L_{00}(K'') = L''_{00}$ als Integralgleichungen für K umdeuten. Demnach hat K für L_{00} die gleiche Bedeutung wie L_{00} für K . Entsprechend (6. 3) ist

$$K + \sum_v W_v'' \cdot B_v' + \sum_u A_u \cdot W_u$$

der allgemeinste lösende Kern für L_{00} .

d. Im Fall $i = i'' = 0$ vereinfachen sich die Darlegungen der Nr. 6 und 7. Alle A_u , B_v und Eigenlösungen sind zu streichen. L und L_0 gehen in das eindeutige L_{00} über. Die beiden L definierenden inhomogenen Integralgleichungen sind jetzt nach (6. 7, 12)

$$(7. 8) \quad L - \int K(t, s) \cdot L(s, t') ds = -K$$

und

$$(7. 9) \quad L - \int L(t, s) \cdot K(s, t') ds = -K.$$

(Diese Gleichungen findet man in der Literatur.) Nach Nr. 7c löst $W = M - \int L \cdot M' dt'$ eindeutig $M = W - \int K \cdot W' dt'$ und umgekehrt. Die hieran anknüpfende Eigenwerttheorie der Integralgleichungen wird in vorliegender Arbeit nicht behandelt (vgl. Hell-Toepl. und Courant).

e. Ergänzung zur Theorie über die Fredholmsche Gleichung, vgl. Nr. 5a. Außer den bisher in Nr. 6 und 7 gemachten Voraussetzungen I bis III der Nr. 1 nehme ich jetzt noch IV hinzu. Ich beweise die am Schluß der Nr. 4g beiläufig angewandte Formel

$$(7. 10) \quad \begin{aligned} & \sum_s 1^{\frac{\epsilon}{n}} \cdot \int L(K^{(n)}; t, s) \cdot K^{(\epsilon)}(s, t') ds \\ &= \sum_s 1^{\frac{\epsilon}{n}} \cdot \int K^{(\epsilon)}(t, s) \cdot L(K^{(n)}; s, t') ds \end{aligned}$$

für $e = 1 \dots (n-1)$ im Fall $i(K^{(n)}) = 0$ und damit nach (4. 4) gleichbedeutend im Fall $i = 0$ und $n = g$.

In (6. 11) für $1^{\frac{e}{n}} \cdot K$ anstatt K , also in

$$(7. 11) \quad L(1^{\frac{e}{n}} \cdot K'')'' = L(1^{\frac{e}{n}} \cdot K)$$

setze ich (4. 18) und

$$\begin{aligned} L(1^{\frac{e}{n}} \cdot K'')'' &= -\Sigma_e 1^{\frac{e}{n}} \cdot K^{(e)} + L(K^{(n)}) + \\ &+ \Sigma_e 1^{\frac{e}{n}} \cdot \int K^{(e)}(t, s) \cdot L(K^{(n)}; s, t') ds \end{aligned}$$

ein. Die letzte Formel entsteht aus (4. 18), wenn man (6. 11) für $K^{(n)}$, das heißt

$$L(K^{(n)})'' = L(K^{(n)})$$

und

$$L(K^{(n)''}; t', s) = L(K^{(n)''}; s, t')' = L(K^{(n)}; s, t')$$

berücksichtigt. Dann folgt aus (7. 11) die Behauptung (7. 10).

Man kann (7. 10) auch aus den beiden gemäß (7. 8, 9) aufzustellenden Integralgleichungen für $L(1^{\frac{e}{n}} \cdot K)$ im Fall $i = 0$ folgern. Denn durch Subtraktion derselben erhält man

$$\int L(1^{\frac{e}{n}} \cdot K; t, s) \cdot K(s, t') ds = \int K(t, s) \cdot L(1^{\frac{e}{n}} \cdot K; s, t') ds,$$

folglich nach (4. 22), daß

$$\int L(K^{(n)}; t, s) \cdot K(s, t') ds = \int K(t, s) \cdot L(K^{(n)}; s, t') ds$$

gilt. Hieraus ergibt sich durch e -malige Anwendung

$$\int L(K^{(n)}; t, s) \cdot K^{(e)}(s, t') ds = \int K^{(e)}(t, s) \cdot L(K^{(n)}; s, t) ds$$

und somit (7. 10).

Schließlich folgt (7. 10) aus (4. 7) und Satz II der Nr. 4b, das heißt aus

$$W(1^{\frac{e}{n}} \cdot K) = T_{n,e}(W(K^{(n)})) = W(K^{(n)}) + \Sigma_e 1^{\frac{e}{n}} \cdot \int K^{(e)} \cdot W(K^{(n)})' dt'$$

für $e = 1 \dots (n-1)$. Man setze hierin

$$W(K^{(n)}) = M - \int L(K^{(n)}) \cdot M' dt'$$

ein und bringe das Ergebnis auf die Form

$$W(1^{\frac{e}{n}} \cdot K) = M - \int L(1^{\frac{e}{n}} \cdot K) \cdot M' dt'.$$

Dann ist

$$L(1^{\frac{c}{n}} \cdot K) = - \sum_s 1^{\frac{c}{n}} \cdot K^{(s)} + L(K^{(n)}) + \\ + \sum_s 1^{\frac{c}{n}} \cdot \int K^{(s)}(t, s) \cdot L(K^{(n)}; s, t') ds,$$

was mit (4. 18) verglichen auf (7. 10) führt.

8. Auflösungstheorie unter Anwendung des Kernabsplittingsverfahrens.

In Nr. 4 bis 7 wurde die Auflösungstheorie für die Integralgleichung (1. 1) mittels Iteration der ganzen Integralgleichung entwickelt. Daneben bringe ich nachfolgend eine solche unter Anwendung des in Nr. 2f angedeuteten Kernabsplittingsverfahrens.

a. Ich beweise zuerst, daß für die *Levische Gleichung* $i = i''$ gilt.

Es sei $Q = Q(t, t')$ ein „kleiner Kern“, das heißt, Q erfülle außer den Voraussetzungen für K noch

$$(8. 1) \quad \int |Q| dt < a \quad \text{und} \quad \int |Q| dt' < a;$$

hierin bedeute a eine Konstante, die $0 < a < 1$ genügt und sonst beliebig ist (Levi Nr. 2 und Hell-Toepl., S. 1379, insbesondere Fußnote 63 daselbst). Nach Voraussetzung VIII der Nr. 1 ist ein Glied der Reihe

$$\Sigma(Q, M) \equiv M + \int Q \cdot M' dt' + \\ + \iint Q(t, s) \cdot Q(s, t') \cdot M' dt' ds + \dots$$

beschränkt. Von diesem Glied an ist wegen (8. 1) die Reihe gleichmäßig und absolut konvergent. Somit gilt die Identität

$$(8. 2) \quad M' = \Sigma(Q, M) - \int Q \cdot \Sigma(Q, M)' dt'.$$

Iteriert man $W - \int Q \cdot W' dt' = 0$ so lange, bis man auf das nach Voraussetzung beschränkte Glied der Reihe $\int Q \cdot W' dt'$, $\iint Q(t, s) \cdot Q(s, t') \cdot W' dt' ds$ usw. kommt, so ergeben c weitere Iterationen gemäß (8. 1), daß $W \prec a^c$ ist; daraus folgt $W = 0$. Somit hat

$$(8. 3) \quad W - \int Q \cdot W' dt' = M$$

den Defekt Null und die eindeutige Lösung

$$(8. 4) \quad W = W(Q) = \Sigma(Q, M).$$

Umgekehrt wird (8. 4) als Gleichung für M eindeutig durch (8. 3) gelöst.

Man kann (nach Levi) in b endlichviele stetige Funktionen $A_c = A_c(t)$ und $B_c = B_c(t)$ so angeben, daß $K - \Sigma_c A_c \cdot B_c' = Q$ ist. Dann folgt aus der

Gleichwertigkeit von (8. 3) mit (8. 4) diejenige der Levischen Gleichung (1. 1) mit

$$(8. 5) \quad W - \int \Sigma_c \Sigma(Q, A_c) \cdot B_c' \cdot W' dt = \Sigma(Q, M).$$

Der Kern von (8. 5) ist stetig, da (nach den Voraussetzungen VI und VII der Nr. 1) jedes Glied der Reihe $\Sigma(Q, A_c)$ stetig, also $\Sigma(Q, A_c)$ eine gleichmäßig konvergente Reihe stetiger Funktionen ist und Σ_c nur aus endlichvielen Gliedern besteht. $\Sigma(Q, M)$ ist von der Art wie M . Demnach gelten für (8. 5) (nach dem ersten Satz der Nr. 2c) die Hauptsätze. Folglich ist (1. 1) unter $i(\Sigma_c \Sigma(Q, A_c)' \cdot B_c)$ notwendigen und hinreichenden Bedingungen für $\Sigma(Q, M)$ lösbar. Ferner folgt aus (8. 2) die Gleichwertigkeit der homogenen Fälle von (1. 1) und (8. 5), somit, wenn man für (8. 5) den zweiten Defektsatz heranzieht,

$$(8. 6) \quad \begin{aligned} i &= i(K) = i(\Sigma_v \Sigma(Q, A_v)' \cdot B_v) \\ &= i(\Sigma_v \Sigma(Q, A_v) \cdot B_v'). \end{aligned}$$

Ebensolches gilt für die zu (1. 1) transponierte Gleichung (6. 8). Jedoch ist die dann an Stelle von (8. 5) tretende Gleichung nicht zu (8. 5) transponiert, also nicht von der Art wie nachfolgende Gleichung (8. 7).

Ich multipliziere (1. 1) mit W_v'' für $v = 1 \dots i''$, integriere und vertausche die Integrationsfolge des Doppelintegrals. Das ist hier zum Unterschied von Kerniterationen auch beim Levischen K erlaubt (vgl. Levi, Anm. I und II der Nr. 3 daselbst). So ergeben sich für (1. 1) die i'' notwendigen Lösungsbedingungen (2. 1), folglich, wegen (8. 2) und wiederum erlaubter Vertauschung der Integrationsfolge,

$$\int \Sigma(Q, M) \cdot (W_v'' - \int Q'' \cdot W_v'' dt) dt = 0.$$

Das sind genau i'' notwendige Bedingungen für $\Sigma(Q, M)$; denn die Funktionen $(W_v'' - \int Q'' \cdot W_v'' dt)$ sind untereinander linear unabhängig, da die zu (8. 3) transponierte Gleichung ebenso wie (8. 3) den Defekt Null hat. Hieraus folgt mit Rücksicht auf (8. 6), daß $i'' \leq i$ ist. Von K'' ausgehend findet man entsprechend $i \leq i''$. Also besteht

$$i = i''.$$

b. Im allgemeinen kann man zur Levischen Gleichung (1. 1) keinen lösenden Kern angeben (vgl. Nr. 2f) und damit auch keine Darstellungen von $W = W(K)$ durch $L(K)$ und $W'' = W(K'')$ durch $L(K'')$, aus denen auf Grund von (6. 11) sich ein Zusammenhang zwischen $W(K)$ und $W(K'')$ ergeben würde. Um einen solchen Zusammenhang auch ohne $L(K)$ nachzuweisen, gehe ich von der zu (1. 1) transponierten Gleichung (6. 8) aus. Eine identische Umformung davon ist

$$(8. 7) \quad (W'' - \int W'''(s) \cdot Q(s, t) ds) - \int \Sigma_c B_c \cdot \Sigma(Q, A_c)' \cdot (\dots)' dt' = M''.$$

Folglich ist

$$W''' - \int Q'' \cdot W''' \, dt' = M''' - \int L(\Sigma_c B_c \cdot \Sigma(Q, A_c)') M''' \, dt'$$

für ein spezielles W'' , wenn ein solches existiert. Ferner ist entsprechend (8.2)

$$\begin{aligned} W''' &= \Sigma(Q'', W'') - \int Q'' \cdot \Sigma(Q'', W'')' \, dt' \\ &= \Sigma(Q'', W'' - \int Q'' \cdot W''' \, dt'). \end{aligned}$$

Also gilt allgemein

$$(8.8) \quad W''' = \Sigma(Q'', M'') - \int \Sigma[Q'', L(\Sigma_c B_c \cdot \Sigma(Q, A_c)')] \cdot M''' \, dt' + \Sigma_c k'_r \cdot W''_r,$$

wenn W'' existiert. Andererseits folgt aus (8.5) allgemein

$$(8.9) \quad W = \Sigma(Q, M) - \int L(\Sigma_c \Sigma(Q, A_c) \cdot B_c') \cdot \Sigma(Q, M)' \, dt' + \Sigma_u k_u \cdot W_u,$$

wenn W existiert. Demnach hängen W und W'' im Existenzfall durch (8.9) und (8.8) zusammen, wobei gemäß (6.11) noch

$$(8.10) \quad L(\Sigma_c B_c \cdot \Sigma(Q, A_c)') = L(\Sigma_c \Sigma(Q, A_c) \cdot B_c')''$$

besteht. Die Bedingungen für M und M'' zur Existenz von W und W'' sind wie bisher (2.1) und (6.9) (wie schon Levi angegeben hat). Denn mit W'' existiert $W''' - \int Q'' \cdot W''' \, dt'$ und umgekehrt. Diese Integraltransformation von W'' löst (8.7) dann und nur dann, wenn M'' orthogonal zu allen Eigenlösungen der zu (8.7) transponierten Gleichung ist. Letztere ist (8.5). Die Eigenlösungen von (8.5) stimmen genau mit denen von (1.1) überein. Diese sind die W_u . Also muß (6.9) gelten und ebenso (2.1).

e. Die als gemeinsamer Sonderfall der Fredholmschen und Levischen Gleichungen angekündigte Integralgleichung (1.1) mit $K = N:R^p$ ist eine Levische Gleichung. Denn N , M und W sind integrierbar und beschränkt (nach den Voraussetzungen IX und X der Nr. 1). Folglich genügt K nach (3.1) den Voraussetzungen V und VI der Nr. 1. Ferner befriedigt K die Voraussetzung VII; denn $R = 0$ ist die einzige Unstetigkeitsstelle des Kernes bei festem t bzw. t' . Schließlich ist die Voraussetzung VIII erfüllt, weil alle Integrale der Voraussetzung I der Nr. 1 nach (3.5) stetig sind.

Ich gebe einen lösenden Kern $L(N:R^p)$ an, indem ich den durch (8.9, 8, 10) bestimmten Zusammenhang zwischen W und W'' spezialisiere. Für $K = N:R^p$ läßt sich Q wesentlich einfacher angeben als für den allgemeinen Levischen Kern (Levi Nr. 2 daselbst). Man nehme zu Q erstens K , solange R kleiner als eine hinreichend klein gewählte Konstante ist, und zweitens die Differenz zwischen dem restlichen stetigen K und einem geeigneten stetigen Annäherungskern endlichen Ranges. (Zur Differenzbildung vgl. Hell-Toepl., Fußnoten 61 und 64. Das dort angegebene Verfahren

läßt sich auf mehrdimensionale Fälle übertragen. Die dabei auftretenden endlichen Sprünge der Glieder des Annäherungskernes hindern nicht die vorliegenden Ausführungen. Man kann sie durch geeignete stetige Übergänge umgehen.) Nach (8.4) ist wegen (3.9) $L(Q) = -\Sigma_v Q^{(v)}$ für $v = 1, 2, \dots$ der lösende Kern im engeren Sinne zu (8.3). $L(Q)$ erfüllt (wegen Nr. 3c) ebenso wie Q die Eigenschaft IX der Nr. 1 wie K . Wählt man m wie in Nr. 3c, so ist $\Sigma_v Q^{(v)}$ vom m -ten Glied an absolut und gleichmäßig konvergent. Also ist $\Sigma_v Q^{(v)}$ ein „kleiner lösender Kern“, wenn entsprechend (8.1) $a < 1:2$ gewählt wird. Denn dann gilt

$$\int |L(Q)| dt \text{ und } \int |L(\dot{Q})| dt' < a: (1-a) < 1.$$

Ich zeige, daß

$$(8.11) \quad L(N:R^p) = L(\Sigma_c(A_c + \int \Sigma_v Q^{(v)} \cdot A_c' dt') \cdot B_c) - \\ - \Sigma_v Q^{(v)} + \int L(\Sigma_c(\dots) \cdot B_c(s)) \cdot \Sigma_v Q^{(v)}(s, t') ds$$

für $v = 1, 2, \dots$ ein lösender Kern zu (1.1), ferner $L(N'':R^p)$ einer zu (6.8) ist und dabei wieder (6.11) gilt. Denn aus (8.9, 10, 8) folgen

$$W = M - \int L(N:R^p) \cdot M' dt' + \Sigma_u k_u \cdot W_u,$$

$$L(\Sigma_c(A_c + \int \Sigma_v Q^{(v)} \cdot A_c' dt') \cdot B_c) = L(\Sigma_c(A_c + \int \Sigma_v Q^{(v)} \cdot A_c' dt') \cdot B_c)''$$

und

$$W'' = M'' - \int L(N:R^p)'' \cdot M''' dt' + \Sigma_v k_v'' \cdot W_v''.$$

Ferner hat $L(N:R^p)$ die allgemeinen Eigenschaften wie $N:R^p$. Beachtet man, daß die Endlichkeit des Defektes und der Lösungssatz auch für die allgemeine Levische Gleichung bestehen, so folgt aus $i = i''$ und (8.11), daß für (1.1) mit $K = N:R^p$ die Hauptsätze gelten. Da ferner auch die Voraussetzungen I, II und III der Nr. 1 erfüllt sind, besteht wieder die in Nr. 6 und 7 entwickelte Theorie für L . Damit habe ich *für den Sonderfall mit $K = N:R^p$ die Gültigkeit der Auflösungstheorie* unter Anwendung des Kernabsplittingsverfahrens nachgewiesen, also neben Nr. 5b auf eine zweite Weise.

Von dem Fredholmschen K läßt sich im allgemeinen nicht die gleichmäßige Konvergenz der Levischen Integrale oder die Möglichkeit der damit zusammenhängenden Kernabsplattung zeigen. Andererseits läßt sich vom Levischen K im allgemeinen nicht die Vertauschbarkeit der Integrationsfolgen und damit nicht die Möglichkeit der Iteration des ganzen K nachweisen. Bedenkt man jedoch, daß das Levische K im erweiterten Sinn uneigentlich singulär, dagegen das Fredholmsche K wie der gemeinsame Sonderfall $K = N:R^p$ im engeren Sinn uneigentlich singulär ist, so ist vom Standpunkt der Auflösungstheorie aus das Levische K der allgemeinere Kern.

II. Wesen und Bildung der Parametrix P für Randwertprobleme zweiter Ordnung im Zweidimensionalen.

Die Stetigkeits- und Differenzierbarkeitsvoraussetzungen von den nachfolgend eingeführten Funktionen lassen sich einschränken, was in vorliegender Arbeit nicht ausgeführt wird.

9. Der Ausgangsbereich für die unabhängigen Veränderlichen.

Ich gebe zuerst den Grundbereich an, für den die Randwertprobleme aufzustellen sind.

a. Die unabhängigen Veränderlichen x und y deute ich als rechtwinklige Koordinaten eines Punktes t einer Ebene. Der Ausgangsbereich b für t sei reell, endlichvielfach zusammenhängend, schlicht und abgeschlossen. Mehrdimensionale Ausgangsbereiche werden erst in den Nrn. 36 und 37 angeführt. Den Rand von b bezeichne ich mit d und den offenen Bereich mit $(b - d)$. Im allgemeinen setze sich d aus endlichvielen, etwa c geschlossenen *Randlinien* zusammen. Für jede derselben bestehe eine *Parameterdarstellung* durch die Bogenlänge s als unabhängige Veränderliche, die ich einheitlich mit

$$x = X = X(s), \quad y = Y = Y(s)$$

bezeichne, und für die ich noch genaue Voraussetzungen angebe. Auf die v -te geschlossene Randlinie, deren Länge ich mit $(s_v - s_{v-1})$ bezeichne, ent falle der s -Bereich $s_{v-1} \leq s \leq s_v$; $v = 1 \dots c$ und $s_0 = 0$.

b. Zuerst nehme ich den Rand d *eckenfrei* an und schreibe $d = d_0$, das $b = b_0$ begrenzt. Dann seien X und Y in jedem Bereich $s_{v-1} \leq s \leq s_v$, eindeutige, viermal stetig differenzierbare Funktionen, wenn die (in Nr. 10a ausführlich angegebene) allgemeine elliptische Differentialgleichung dem aufzustellenden Randwertproblem zugrunde liegt. Wird jedoch diese Differentialgleichung auf ihre Normalform (vgl. Nr. 10a) beschränkt, so genügt die dreimalige stetige Differenzierbarkeit. X habe der Geschlossenheit der Randlinien wegen für $(s_{v-1} + 0)$ und $(s_v - 0)$ die gleichen Werte und ist im Bereich $0 \leq s \leq s_c$ im Fall $c \geq 2$ doppeldeutig für $s = s_v$; $v = 1 \dots (c - 1)$. Das selbe gilt für Y , ferner für jede der genannten Ableitungen von X und Y .

Zu jedem s gehört eindeutig eine Normale durch (X, Y) . Es sollen die positive s - und positive Normalenrichtung wie die positive x - und positive y -Richtung der orientierten Ausgangsebene aufeinander folgen. Dabei wähle ich den bei wachsendem s erfolgenden Durchlaufungssinn der Randlinien von vornherein so, daß die negative Normalenrichtung bei stetiger Abhängigkeit von s stets in das Äußere eines hinreichend kleinen, bei (X, Y) liegenden, zusammenhängenden Teiles von b_0 zeigt. [Bedarfsweise ist in $X(s)$, $Y(s)$ das $(+s)$ durch $(-s)$ zu ersetzen.] Jeder Punkt $t = (x, y)$ der zu s gehörenden

Normalen hat einen mit Vorzeichen versehenen Abstand n von (X, Y) . Dabei kann hier t innerhalb und außerhalb b liegen. Also gehört zu jedem Wertepaar (s, n) eindeutig

$$x = x(s, n) = X - Y_s \cdot n,$$

$$y = y(s, n) = Y + X_s \cdot n$$

für beliebige n und solche s , die $s_{v-1} \leq s \leq s_v$ genügen; $v = 1 \dots c$. Umgekehrt läßt sich eine Kette von endlichvielen abgeschlossenen Bereichen für (s, n) der Art

$$s_{v-1} + \frac{s_v - s_{v-1}}{f_v} \cdot (u-1) \\ \leq s \leq s_{v-1} + \frac{s_v - s_{v-1}}{f_v} \cdot u \quad \text{und} \quad 0 \leq n \leq n_{uv}$$

so angeben, daß in einem beliebigen Bereich der Kette jedem (x, y) eindeutig ein (s, n) entspricht; hierin seien die f_v genügend große, ganze (endliche) Zahlen, n_{uv} hinreichend kleine, positive Zahlen, $u = 1 \dots f_v$, $v = 1 \dots c$. (Die Bereiche können über b_0 hinausgreifen.) In jedem dieser Bereiche ist $n(x, y)$ eine stetige Funktion mit stetigen vierten bzw. dritten Ableitungen nach x und y , wenn die allgemeine elliptische Differentialgleichung bzw. deren Normalform vorliegt. Entsprechendes gilt für $s(x, y)$ mit stetigen dritten bzw. zweiten Ableitungen. Denn es ist

$$n_x = -Y_s, \quad n_y = +X_s, \\ s_x = \frac{X_s}{1 - n \cdot (X_s \cdot Y_{ss} - Y_s \cdot X_{ss})}, \\ s_y = \frac{Y_s}{1 - n \cdot (X_s \cdot Y_{ss} - Y_s \cdot X_{ss})} \quad \text{usw}$$

Alle diese Funktionen bleiben stetig, wenn t über eine Grenznormale in einen Nachbarbereich der angegebenen Kette von Bereichen übergeht, mit Ausnahme der Doppeldeutigkeit von s für t auf $s = s_v$; $v = 0 \dots c$.

Es sei — und das ist eine wesentliche Voraussetzung für b_0 — eine positive Zahl $m < n_{uv}$ so angebbar, daß die eindeutig umkehrbare Zuordnung zwischen (x, y) und (s, n) in dem System offener „Randstreifen“ $0 < n \leq m$, $0 \leq s \leq s_v$ mit Ausnahme der genannten Doppeldeutigkeit für s gelte, und daß diese Streifen ganz zu b_0 gehören. Der Ausschluß von $n = 0$ bedeutet, daß d_0 Doppelstücke haben darf. $n(x, y)$ ist in $0 \leq n \leq m$, $0 \leq s \leq s_v$ viermal bzw. dreimal stetig differenzierbar mit Ausnahme der Doppeldeutigkeit der Ableitungen in den Doppelstücken von d_0 . Man kann somit b_0 als einen endlichvielfach zusammenhängenden Bereich ohne Abschnürungen bezeichnen, dessen Randlinien sich nicht selbst überschneiden und sich nicht wechselseitig treffen. Jede einzelne Randlinie für sich hat im allgemeinen Doppel-

stücke und erfüllt gewisse Differentialeigenschaften, die über die Stetigkeit der Krümmung hinausgehen.

e. Ferner habe der Rand d gewisse Ecken. Dann sollen X und Y die gleichen Voraussetzungen wie für d_0 erfüllen, ausgenommen in endlichvielen Ecken mit $s = s_{vw}$ für $w = 1 \dots p$. Hierbei wähle man von vornherein $s_{vw} \neq s_v$ für $v = 0 \dots c$. Ecken mit untereinander verschiedenen s_{vw} werden auch dann als verschieden angesehen, wenn sie in den gleichen mehrfachen Randpunkt fallen. In den Ecken soll mindestens eine der genannten Ableitungen von X und Y einen endlichen Sprung haben. Der Winkel, gezählt von der positiven Richtung der Grenztangente für $(s_{vw} + 0)$ an im positiven Drehsinn der x - y -Ebene bis erstmals die negative Richtung der Grenztangente für $(s_{vw} - 0)$ erreicht wird, heiße Eckenwinkel e_w .

I. Es seien alle Eckenwinkel e_w untereinander gleich, etwa $e_w = e$, und es falle keine Ecke auf einen der möglichen Doppelpunkte des Randes d . Zu jeder einzelnen Randlinie von d sollen sich genau zwei geschlossene Hilfslinien angeben lassen, die

1. die allgemeinen Eigenschaften jeder Einzellinie eines d_0 haben,
2. $(b - d)$ meiden,
3. beide zusammengenommen diese Randlinie von d restlos (einfach und erlaubterweise auch doppelt) überdecken.

Dann muß

$$(9.1) \quad 0 \leq e \leq \pi$$

sein. Die beiden Systeme von je einer Hilfslinie für jede geschlossene Randlinie bezeichne ich mit d_0 und d_{00} , die von ihnen begrenzten Bereiche mit b_0 und b_{00} und die dazugehörigen n mit n_0 und n_{00} .

II. Es seien entweder mindestens zwei der Eckenwinkel e_w untereinander verschieden, oder es beteilige sich eine endliche Anzahl von Ecken mit untereinander verschiedenen zugehörigen Bogenlängen s_{vw} an einem mehrfachen Randpunkt, oder es sei beides der Fall. Unter Verzicht auf allgemeinere Möglichkeiten und ohne zu untersuchen, inwieweit die einzelnen Voraussetzungen auseinander hervorgehen, nehme ich die Existenz nachfolgender Bereiche b_u für $u = 1 \dots q$ an; hierin bedeute q eine endliche Zahl.

Die zu $s = s_{vw}$ gehörende Ecke trenne ich durch einen zu b gehörenden Kreisausschnitt $K(s_{vw}, a)$ mit dem Halbmesser a und dieser Ecke als Mittelpunkt von b ab. Dabei sei $K(s_{vw}, a)$ einfach zusammenhängend und enthalte keine mehrfachen Punkte des Randes d . Ferner verlaufe das $K(s_{vw}, a)$ abgrenzende Kreisbogenstück mit Ausnahme von nur zwei Punkten in $(b - d)$ und berühre kein $K(s_{vw}, a)$ für $v \neq w$. Die beiden ausgenommenen Punkte sind die auf d gelegenen Endpunkte. Solches gelte für jedes a , das etwa

$0 < a \leq a_0$ genügt. Insbesondere werde jetzt für a irgendein Wert gewählt, der kleiner als $a_0 : 6$ ist.

b_1 sei ein Teilbereich von b von der in Nr. 9b angegebenen Art wie b_0 oder von der vorstehend unter I angegebenen Art. b_1 enthalte $b - \sum_w K(s_w, 2a)$ für $w = 1 \dots p$ und dazu irgendeines der $K(s_w, 2a)$ oder mehrere $K(s_w, 2a)$ mit untereinander gleichen e_w , sofern jede zugehörige Ecke keinen mehrfachen Randpunkt von b_1 verursacht. Diese Ecken dürfen jedoch an einem mehrfachen Randpunkt von b teilnehmen. Es soll aber dem b_1 nur ein einziger solcher Wert e_w zugeordnet werden (sofern das überhaupt möglich ist). b_1 berühre kein $K(s_w, a)$ mit den übrigen Ecken.

b_2 sei ein Teilbereich von b von der unter I angegebenen Art oder die Zusammenfassung einer endlichen Anzahl solcher getrenntliegender Teilbereiche. b_2 enthalte irgendein $K(s_w, 5a)$ mit einer der übriggebliebenen Ecken. Mindestens ein solches $K(s_w, 5a)$ ist vorhanden. Oder b_2 enthalte mehrere solcher $K(s_w, 5a)$ mit untereinander gleichen e_w , sofern keine zwei der zugehörigen Ecken in den gleichen Randpunkt von b_2 fallen. Diese Ecken dürfen jedoch an einer mehrfachen Ecke des Randes von b teilnehmen. Auch dem b_2 soll nur ein einziger Wert e_w zugeordnet werden, welcher gleich dem in b_1 auftretenden sein könnte (wenn e_w zu einem mehrfachen Randpunkt gehört). Ferner sei b_2 ein Teilbereich des zur gleichen Ecke gehörenden $K(s_w, 6a)$ bzw. die Summe von Teilbereichen der zu den gleichen Ecken gehörenden getrenntliegenden $K(s_w, 6a)$. b_2 deckt sich in Ringstücken der Art $K(s_w, 6a) - K(s_w, a)$ teilweise mit b_1 .

Sind noch nicht alle Ecken den b_1 und b_2 zugewiesen, so fahre man mit der Bildung von b_3, b_4 usw. entsprechend b_2 fort. Nach etwa q Schritten sind die p Ecken verteilt. q ist eine endliche Zahl größer als eins. (Es kann $q > p$ sein; beispielsweise ist $q = p + 1$, wenn b nur eine Ecke hat und diese in einem Doppelpunkt des Randes liegt.) Ich nummeriere (aus formalen Gründen) die s_w so um, daß die zugehörigen e_w für $w = 1 \dots q$ gerade zu b_w gehören. Ist b_1 von der Art b_0 , so lasse ich die Bezeichnung e_1 ausfallen. Die restlichen Ecken, deren zugehörige e_w für $w \geq q + 1$ mit gewissen e_w für w aus der Reihe $1 \dots q$ übereinstimmen, können nach Belieben verteilt werden.

Liegt der Punkt t in b , so liegt er entweder nur in einem einzigen b_u für $u \geq 1$, oder in dem gemeinsamen Teil b_{1u} von b_1 und einem b_u für $u > 1$. b_{1u} oder gegebenenfalls jeder zusammenhängende Teil von b_{1u} ist in einem der Ringstücke $K(s_w, 6a) - K(s_w, a)$ enthalten. Ich kann und will b_{1u} von der Art wie b_0 wählen. Insbesondere sollen die für b_u zu bildenden $d_0 = d_{0u}$ und $d_{00} = d_{00u}$ für $u > 1$ innerhalb $K(s_w, 6a) - K(s_w, 3a)$, ferner für $u = 1$ innerhalb $K(s_w, 4a) - K(s_w, a)$ genau mit dem für b_{1u} zu bildenden eckenfreien d_{1u} zusammenfallen. Dabei ist im Fall $b_1 = b_0$ das d_{001} zu streichen.

Allgemein müssen die d_{0u} und d_{00u} die $(b_u - d_u)$ meiden, aber nicht die $(b_v - d_v)$ für $v \neq u$. Nach (9. 1) muß stets $0 \leq e_u \leq \pi$ sein.

d. Zur genauen Durchführung der Auflösungstheorie setze ich alle vorkommenden $e_w = \pi : g_w$; hierin sind die g_w irgendwelche ganze, positive Zahlen größer als eins. Das wird in Nr. 39 näher begründet. Allgemeinere e werden bei der Abgrenzung der Gültigkeit der Parametrixmethode (in Nr. 38 und 40) angeführt. Die zulässigen eckigen b und d bezeichne ich für $e_w = \pi : g_w$ mit $b = b(\pi : g_1, g_2, \dots, g_p)$ und $d = d(\pi : g_1, g_2, \dots, g_p)$. Dann kann ich die Hilfsbereiche b_u mit

$$b_1 = b_0 \quad \text{oder} \quad b_1 = b(\pi : g_1, g_1, \dots)$$

und ferner

$$b_u = b(\pi : g_u, g_u, \dots) \quad \text{für} \quad u = 2 \dots q$$

bezeichnen. Es kann auch ein $b(\pi : g_u)$ an Stelle von $b(\pi : g_u, g_u, \dots)$ treten. Die vier geradlinig begrenzten $b(\pi : g_1, g_2, \dots, g_p)$ werden in Nr. 15a angegeben.

10. Angabe der zu lösenden Randwertprobleme.

a. Vorgelegt sei die allgemeine lineare partielle *Differentialgleichung* zweiter Ordnung vom elliptischen Typus (Lichtenstein, Diff.gl., S. 1294):

$$(10. 1) \quad EU = A \cdot U_{xx} + 2B \cdot U_{xy} + C \cdot U_{yy} + A_0 \cdot U_x + B_0 \cdot U_y + C_0 \cdot U = F \text{ in } (b - d)$$

mit $A \cdot C - B^2 = 1$; ferner sei A etwa positiv. $U = U(t) = U(x, y)$ sei gesucht und die übrigen Funktionen seien vorgegeben; ich nehme dabei folgendes an.

$$(10. 2) \quad F = F(t) = F(x, y) \text{ sei stetig in } b \text{ und stetig differenzierbar in } (b - d); \text{ ferner seien}$$

$$F_x \text{ und } F_y \lesssim \log n \text{ für } b = b_0 \text{ und} \\ \lesssim \Sigma_v C_v \cdot \log n_v \text{ mit } v = 1 \dots q \text{ für } b = b(\pi : g_1, g_2, \dots, g_p).$$

(Das Zeichen \lesssim wurde am Schluß der Vorbemerkung zu Nr. 3 eingeführt.) Zur Abschätzung durch $\log n$ und $\log n_v$ setze ich voraus:

I. t liege im Randstreifensystem $0 \leq n \leq m$, $0 \leq s \leq s_e$ längs d_0 und in einem solchen Streifensystem längs des Randes d_{0v} oder d_{00v} von irgendeinem der (in Nr. 9c angegebenen) Teilbereiche b_v von $b(\pi : g_1, g_2, \dots, g_p)$. Der Rand von b_1 kann auch von der Art wie d_0 sein.

II. n_v sei die kleinere der in den genannten Randstreifensystemen erklärten Funktionen n_{0v} und n_{00v} .

III. $C_v = C_v(t)$ seien etwa stetige Funktionen in b ($\tau: g_1, g_2, \dots, g_p$). Insbesondere gelte:

1. $C_v = 0$ in b außerhalb b_v .
2. $C_1 = 1$ in b_1 außerhalb aller b_{1u} für $u > 1$ und $C_v = 1$ in b_v außerhalb von b_{1v} für $v > 1$.
3. $C_v \cong n_{1v} = n_{0v} = n_{00v}$ in einem Randstreifen längs $d_{1v} = d_{0v} = d_{00v}$, sofern dieser entweder für $v = 1$ innerhalb von jedem $K(s_w, 3a)$, das zu einem b_u für $u > 1$ gehört, oder für $v > 1$ außerhalb der zu b_v gehörenden $K(s_w, 4a)$ liegt.

Im übrigen seien die C_v beliebig. Man könnte wegen der Eigenschaft III, 1 die etwa getrennt liegenden Teile von b_u für $u \geq 2$ formal zu einem zusammenhängenden b_u durch eine Brücke so verbinden, daß b_u alle allgemeinen Eigenschaften von b ($\tau: g, g, \dots$) hat.

F habe also diese unter (10. 2) angegebenen Eigenschaften. Ebenso sei C_0 von der Art wie F . A_0 und B_0 seien stetig differenzierbar in b mit ersten Ableitungen von der Art wie F . A, B und C seien zweimal stetig differenzierbar in b mit zweiten Ableitungen von der Art wie F .

- (10. 3) U sei stetig in b und zweimal stetig differenzierbar in $(b - d)$; U_x und U_y seien beschränkt in b ; EU sei integrierbar in b .
- (10. 4) Es existiere $E_0U = \lim EU$ für $(x, y) \rightarrow (X, Y)$. Setze ich ferner $E_0U = EU$ in $(b - d)$, so ist E_0U stetig in b .
- (10. 5) E_0U sei von der Art wie F in b .

An Stelle von E tritt mitunter einer der nachfolgenden Differentialausdrücke. Die Normalform von EU ist

$$U_{xx} + U_{yy} + A_0 \cdot U_x + B_0 \cdot U_y + C_0 \cdot U = \Delta U + TU,$$

wenn ich

$$\Delta U = U_{xx} + U_{yy} \quad \text{und} \quad TU = A_0 \cdot U_x + B_0 \cdot U_y + C_0 \cdot U$$

setze. Der zu EU adjungierte Ausdruck ist

$$\begin{aligned} \tilde{E}U &= \tilde{A} \cdot U_{xx} + 2 \cdot \tilde{B} \cdot U_{xy} + \tilde{C} \cdot U_{yy} + \tilde{A}_0 \cdot U_x + \\ &\quad + \tilde{B}_0 \cdot U_y + \tilde{C}_0 \cdot U = (A \cdot U)_{xx} + 2 \cdot (B \cdot U)_{xy} + \\ &\quad + (C \cdot U)_{yy} - (A_0 \cdot U)_x - (B_0 \cdot U)_y + C_0 \cdot U. \end{aligned}$$

Hierin erfüllen

$$\begin{aligned} \tilde{A} &= A, \quad \tilde{B} = B, \quad \tilde{C} = C, \\ \tilde{A}_0 &= 2 \cdot A_x + 2 \cdot B_y - A_0, \\ \tilde{B}_0 &= 2 \cdot B_x + 2 \cdot C_y - B_0, \\ \tilde{C}_0 &= A_{xx} + 2 \cdot B_{xy} + C_{yy} - A_{0x} - B_{0y} + C_0 \end{aligned}$$

die oben für A, B, C, A_0, B_0, C_0 angegebenen Eigenschaften. ΔU ist zu sich selbst adjungiert. Setzt man $TU = \bar{A}_0 \cdot U_x + \bar{B}_0 \cdot U_y + \bar{C}_0 \cdot U$, so ist $(\Delta U + \bar{T}U)$ die adjungierte Normalform. Die zu $\bar{E}U$ und $(\Delta U + \bar{T}U)$ adjungierten Ausdrücke sind wieder EU und $(\Delta U + TU)$.

$t' = (x', y')$ sei ein Parameterpunkt. Ich schreibe

$$U' = U(t'), \quad A' = A(t') \text{ usw.},$$

ferner

$$E'U' = A' \cdot U'_{x'x'} + 2 \cdot B' \cdot U'_{x'y'} + \dots$$

und entsprechend $\Delta'U', T'U'$ und die adjungierten Fälle. Für eine Funktion

$$S = S(t, t') = S(x, y; x', y')$$

setze ich

$$ES = A \cdot S_{xx} + 2 \cdot B \cdot S_{xy} + \dots,$$

$$E'S = A' \cdot S_{x'x'} + 2 \cdot B' \cdot S_{x'y'} + \dots \text{ usw.}$$

Schließlich führe ich

$$DS = A' \cdot S_{xx} + 2 \cdot B' \cdot S_{xy} + C' \cdot S_{yy}$$

und

$$D'S = A \cdot S_{x'x'} + 2 \cdot B \cdot S_{x'y'} + C \cdot S_{y'y'}$$

ein.

Jeder der anschließend E genannten Differentialausdrücke ist entsprechend E_0 zu bilden, wenn man ihn mit dem Index Null versieht.

b. Die Randbedingung sei von der „ersten Art“, das heißt

$$(10.6) \quad U = 0 \quad \text{auf } d.$$

Hierauf läßt sich der allgemeinere Fall $U = M$ auf d zurückführen, wenn $M = M(t)$ in b irgendeine Funktion ist, die (10.3, 4, 5) für M anstatt U erfüllt. M braucht jedoch keiner Differentialgleichung zu genügen (Courant, Bd. 1, S. 222).

c. Die im nachfolgenden Abschnitt IV ausführlich gelösten *eigentlichen* linearen Randwertprobleme erster Art und zweiter Ordnung sind zwei Hauptfälle von (10.1, 6), nämlich erstens

$$(10.7) \quad EU = F \quad \text{in } (b_0 - d_0)$$

und

$$U = 0 \quad \text{auf } d_0$$

$$\text{mit } A \cdot C - B^2 = 1 \quad \text{und } A > 0$$

und zweitens

$$(10.8) \quad \Delta U + TU = F$$

$$\text{in } b(\pi: g_1, g_2, \dots, g_p) - d(\dots)$$

und

$$U = 0 \quad \text{auf } d(\pi: g_1, g_2, \dots, g_p).$$

Die Differentialgleichungen lassen sich auch in der Form

$$E_0 U = F \quad \text{in } b_0$$

und

$$\Delta_0 U + T_0 U = F \quad \text{in } b(\pi: g_1, g_2, \dots, g_r)$$

als Hauptfälle von

$$E_0 U = F \quad \text{in } b$$

schreiben.

„Eigentliches Randwertproblem“ ist in doppeltem Sinn zu verstehen:

1. Die zur Differentialgleichung tretenden Bedingungen sind Randbedingungen und keine sogenannten Anfangswert-, Sprung-, Unstetigkeitsflächen- oder sonstige Zusatzbedingungen. Demnach ist die Differentialgleichung *elliptisch* unter der Annahme der Existenz von eindeutig bestimmten und von den vorgegebenen Daten stetig abhängenden Lösungen, welche die angegebenen allgemeinen Voraussetzungen erfüllen (Courant, Bd. 2, S. 171 bis 179).

2. Der Ausgangsbereich b hat wirklich einen Rand d , ist also kein in sich geschlossener Bereich wie etwa die (in Nr. 12a erwähnte) Kugeloberfläche in Hilberts Beispiel.

Für die beiden eigentlichen Randwertprobleme (10. 7, 8) wird eine Auflösungstheorie in Parallele und im Zusammenhang mit der (im Abschnitt I angegebenen) Auflösungstheorie für Integralgleichungen entwickelt. Die Nr. 2b entsprechenden Hauptsätze findet man in Nr. 27a und b. Ferner wird die Nr. 6 und 7 entsprechende Theorie der allgemeinsten Greenschen Funktion G und kennzeichnender Sonderfälle derselben in Nr. 29 und 30 gebracht. Dabei ergibt sich für jede Lösung U , die (10. 3, 4, 5, 6) genügt, als Verschärfung von (10. 3) und (10. 6):

(10. 9) U ist stetig differenzierbar in b mit ersten Ableitungen von der Art wie F ,

das heißt, U ist wie A_0 und B_0 beschaffen;

$$(10. 10) \quad \begin{aligned} U &\lesssim n \quad \text{für } b = b_0 \quad \text{und} \\ U &\lesssim \Sigma_v C_v \cdot n_v \quad \text{für } b = b(\pi: g_1, g_2, \dots, g_r), \end{aligned}$$

wenn t , n_v und C_v die anschließend (10. 2) eingeführten Voraussetzungen I, II und III erfüllen. Weiter begründe ich in Nr. 33 bis 37, daß die angekündigte Auflösungstheorie auch für verwandte Randwertprobleme höherer Ordnung und mit mehrdimensionalen Ausgangsbereichen entwickelt werden kann. Schließlich werden in Nr. 38 bis 40 die Grenzen der Beweismethode und gewisser Abänderungen derselben im Hinblick auf allgemeinere Probleme fest-

gestellt. Auf den Zusammenhang der Randwertprobleme mit der mathematischen Physik wird in dieser Arbeit nicht eingegangen (vgl. Courant und Frank-Mises).

11. Auflösungsmethoden mittels Greenscher Funktion und Grundlösung.

In der Literatur löst man das Problem (10. 1, 6) meistens durch die nachfolgend erwähnten Integralgleichungsmethoden. Andererseits gibt es Beweise der Hauptsätze auf Grund der Variationsrechnung (Courant, Bd. 2, Kap. VII, insbesondere S. 473) bei Beschränkung auf den zu sich selbst adjungierten Fall $E = \tilde{E}$ (S. 284 daselbst) und mittels zahlreicher Einzelverfahren (vgl. Lichtenstein, Diff.gl., z. B. für hinreichend kleine Ausgangsbereiche im Normalfall: S. 1280–1281, ferner Kornsches Verfahren ohne Transformation des allg. ellipt. Differentialausdrucks auf den Normalfall: S. 1296–1297).

a. Unter dem Defekt r bzw. \tilde{r} des ursprünglichen bzw. adjungierten Randwertproblems (10. 1, 6) verstehe ich die Anzahl der untereinander linear unabhängigen Lösungen des homogenen Problems mit $F = 0$. Man versucht, die sogenannte *Greensche Funktion* $G = G(t, t') = G(x, y; x', y')$ im Fall $r = 0$ mit folgenden Eigenschaften zu bilden.

I. G sei zweimal stetig differenzierbar nach $(x, y; x', y')$ für t und t' in b , solange $R > 0$ ist. Ferner sei $G \lesssim \log R$ und die ersten Ableitungen von G seien $\lesssim 1:R$. Hierin bestimme ich

$$R = R(t, t') \geq 0 \quad \text{und} \quad R^2 = (x - x')^2 + (y - y')^2.$$

I. G sei für $R = 0$ in solcher Weise singulär, daß die sogenannten Greenschen Formeln

$$(11.1) \quad \int G \cdot E' U' dt' = \int \tilde{E}' G \cdot U' dt' - U \quad \text{für } t \text{ in } b$$

und

$$(11.2) \quad E \int G \cdot F' dt' = \int E G \cdot F' dt' - F \quad \text{für } t \text{ in } (b - d)$$

bestehen, wenn man von den Integralen eine kleine Kreisfläche um t mit dem Halbmesser a ausschließt und dann zum Grenzwert für $a = 0$ übergeht.

III. Es sei

$$EG = 0 \quad \text{für } t \text{ in } (b - d)$$

und

$$G = 0 \quad \text{für } t \text{ auf } d, \text{ wenn } t' \text{ in } b \text{ liegt,}$$

ferner

$$\tilde{E}' G = 0 \quad \text{für } t' \text{ in } (b - d)$$

und

$$G = 0 \quad \text{für } t' \text{ auf } d, \text{ wenn } t \text{ in } b \text{ liegt;}$$

beidemale soll $R > 0$ sein.

IV. $\int G \cdot F' dt'$ habe die in (10. 3, 4, 5) angegebenen allgemeinen Eigenschaften von U .

Aus I bis IV folgt, daß

$$(11. 3) \quad U = - \int G \cdot F' dt' \quad \text{für } r = 0$$

die eindeutige Lösung von (10. 1, 6) ist, wenn G existiert. Die Voraussetzungen dafür, daß $r = 0$ eintritt, werden in dieser Arbeit nicht erörtert. (Vgl. die Eindeigkeitssätze in Sommerfeld, Nr. 4 daselbst, weiter Lichtenstein, Diff.gl., S. 1283 und 1297–1298, ferner Courant, Bd. 2, S. 275. — Betreffs $r = 0$ für das $\Delta\Delta$ -Problem vgl. Frank-Mises Teil I, S. 640–641. Der dort angegebene Beweis für ein viermal stetig differenzierbares U und einfach zusammenhängende b läßt sich auf das U vorliegender Arbeit dadurch übertragen, daß man zuerst $0 \leq n \leq k$ ausschließt und dann $k \rightarrow 0$ bildet; vgl. meine Nrn. 33 und 34. Der Beweis läßt sich auch für mehrfach zusammenhängende b durchführen.)

Hat jedoch (10. 1, 6) den Defekt $r > 0$, so versuche man irgendeine Funktion $c = c(t)$ von der Art wie F in (10. 2) so zu bestimmen, daß $EU + c \cdot U = 0$ mit der Randbedingung (10. 6) den Defekt Null hat. c kann auch eine Konstante sein. Kennt man zu letzterem Problem die Greensche Funktion G_c , so gilt

$$(11. 4) \quad U + \int G_c \cdot c' \cdot U' dt' = - \int G_c \cdot F' dt'$$

für die Lösung von (10. 1, 6). Diese Integralgleichung gehört zu der in Nr. 1 d angegebenen und ist nach Abschnitt I zu lösen. Aus ihrer allgemeinen Lösung ist auf die Lösung von (10. 1, 6) zu schließen.

b. Die G -Methode bei eckigen Ausgangsbereichen geht von bekannten G für Δ und für gewisse b_0 aus, die der Potentialtheorie entnommen werden. (Betreffs G und den Greenschen Formeln vgl. z. B. Courant und Lichtenstein, Diff.gl.; daselbst weitere Literaturhinweise. Insbesondere für den Fall Δ bei zweidimensionalem bzw. drei- oder mehrdimensionalem b vgl. Courant, Bd. 1, S. 285, bzw. Bd. 1, S. 291 oder Bd. 2, S. 239. — Betreffs $\Delta\Delta$ -Problem im Zweidimensionalen vgl. Courant, Bd. 1, S. 291.)

I. Man transformiere E auf die Normalform $(\Delta + T)$ bei festbleibendem b . Solches in jedem Punkt von b zu erreichen, beruht auf der Lösbarkeit gewisser Differentialgleichungen. (Vgl. Courant, Bd. 2, S. 123–128, und Lichtenstein, Diff.gl., S. 1294–1295.)

II. Man bilde $(b - d)$ auf ein $(b_0 - d_0)$ konform ab, wobei $(\Delta + T)$ wieder in eine Normalform übergeht. Dabei sei G für Δ und das eckenfreie b_0 bekannt. (Vgl. Lichtenstein, Diff.gl., S. 1283. — Betreffs Zusammenhang von Greenscher Funktion und konformer Abbildung vgl. Courant, Bd. 1, S. 297–298, ferner Frank-Mises, 1. Teil, Kap. XVI von Löwner.)

III. Das Problem mit $(\Delta + T)$ und b_0 wird gemäß Nr. 11a gelöst. Auf Grund von (11. 1) ist

$$(11. 5) \quad U - \int \tilde{T} G \cdot U' dt' = - \int G \cdot F' dt'.$$

Ferner ist auf Grund von (11. 2), wenn man U entsprechend (11. 3) in der Form

$$(11. 6) \quad U = - \int G \cdot Z' dt'$$

und $Z = Z(t)$ als eine Funktion der Art wie F annimmt,

$$(11. 7) \quad Z - \int T G \cdot Z' dt' = F.$$

Beide Integralgleichungen (11. 5, 7) sind nun mit den in Nr. 1 angegebenen verwandt und ähnlich wie im Abschnitt I zu lösen. Aus ihrer allgemeinen Lösung ist bei (11. 5) unmittelbar und bei (11. 7) mittels (11. 6) auf die Lösung von (10. 1, 6) zu schließen. Kennt man jedoch eine „reduzierte Greensche Funktion“, das heißt ein G für E und b , wobei A_0 , B_0 und C_0 durch andere Funktionen der gleichen Art ersetzt sind (beispielsweise durch Null), so kann man (11. 5, 7) unmittelbar ohne vorstehende I und II aufstellen. [Betreffs (11. 5) vgl. Hilbert, Kap. IX und Lichtenstein, Diff.gl., S. 1281–1282 und Fußnoten 8 und 22 daselbst. Riemanns klass. Kennzeichnung dieser Zurückführung in den Gött. Nachr. von 1860 = math. Werke, II. Aufl., S. 170–171. Vgl. auch Hell., Nr. 21 und Sommerfeld. — Betreffs (11. 6, 7) vgl. Courant, Bd. 2, S. 282–284, bei hinreichend kleinem b .]

G für $(\Delta + T)$ und b_0 ist als eine spezielle Lösung des Randwertproblems wieder mittels III angebbbar (Lichtenstein, Diff.gl., S. 1287–1288).

IV. G für $(\Delta + T)$ und eckige b läßt sich als Grenzwert einer unendlichen Folge von G für $(\Delta + T)$ und b_0 bilden. Diese b_0 konvergieren in gewisser Weise nach b (vgl. Lichtenstein, Diff.gl., S. 1291. Eine ähnliche Methode für Δ findet man in Courant, Bd. 2, S. 263–269).

Durch vorstehende G -Methode werden die Hauptsätze bewiesen (vgl. z. B. Lichtenstein, Diff.gl., S. 1282 nebst Fußnote 8, S. 1289 und 1294–1295). Es wird angegeben, daß mit $r = 0$ auch $\tilde{r} = 0$ ist (vgl. Lichtenstein, Diff.gl., S. 1289 nebst Fußnote 24). $r = \tilde{r}$ stellt Hilbert für sein Beispiel mittels seiner nachfolgend in Nr. 12a angedeuteten Parametrixmethode fest. Die Einschränkung der Differentialgleichung auf $(b - d)$ findet sich vor allem in Lichtenstein, Diff.gl. (S. 1279–1281) und Courant, Bd. 2 (S. 223–224, 239). Lichtenstein bzw. Courant geht von einem offenen Gebiet aus, das er mit T bzw. G bezeichnet. In der übrigen Literatur wird vorwiegend die Differentialgleichung im ganzen b aufgestellt. U wird in der Literatur als zweimal stetig differenzierbar bis auf endliche Sprünge in den zweiten Ableitungen vorausgesetzt; jedoch werden die weiteren Annahmen nicht erwähnt.

Wesentlich unvollständigere und speziellere Aussagen macht man bei den Erweiterungen auf Randwertprobleme mit mehrdimensionalem b und auf Probleme höherer Ordnung. Beispielsweise löst man das $\Delta\Delta$ -Problem im Zweidimensionalen durch Angabe des G für $\Delta\Delta$, dessen Existenz auf derjenigen der Lösung zweier simultaner Integralgleichungen fußt. (Frank-Mises, Teil I, Kap. XIX, 2 + 3. — Ausgiebig untersuchte Sonderfälle des $\Delta\Delta$ -Problems in Courant als Plattenschwingungsprobleme, Sternberg, S. 1147 und Sommerfeld, S. 568. Weitere Hinweise in Lichtenstein, Diff.gl., S. 1300 und Fußnoten 22, 35, 38.)

Diese Ergebnisse werden in den Abschnitten IV und V der vorliegenden Arbeit systematisch zusammengestellt und vervollständigt, so wie ich es für die Randwertprobleme (10. 7, 8) in Nr. 10c genauer ankündigte. Das geschieht auf wesentlich elementarere Weise durch die Parametrixmethode, worauf ich in Nr. 12b zurückkomme.

e. Mit obiger G -Methode ist die *Grundlösungsmethode bei eckenfreien Ausgangsbereichen* b_0 verwandt.

I. Man ersetze G in (11. 6, 7) durch $\log Q^2$; hierin bestimme ich $Q = Q(t, t') \geq 0$, $Q^2 = C \cdot (x - x')^2 - 2 \cdot B \cdot (x - x') \cdot (y - y') + A \cdot (y - y')^2$ für $A \cdot C - B^2 = 1$ und $A > 0$, somit $Q = R$ im Fall $E = \Delta$. $\log Q^2$ ist die „reduzierte Grundlösung“. Das ist die bis auf Konstanten einzige Funktion, die wie in (16. 2)

$$D'S = A \cdot S_{x'x'} + 2 \cdot B \cdot S_{x'y'} + C \cdot S_{y'y'} = 0$$

für $R > 0$ löst, wenn man von $S = S(t, t')$ die Eigenschaften I und II von G in Nr. 11a voraussetzt. S genügt im allgemeinen nicht wie G der Randbedingung in III der Nr. 11a. Aus den möglichen Lösungen von (10. 1) bilde man durch ein Grenzwertverfahren eine „vollständige Grundlösung“, das heißt eine Funktion S , die $ES = 0$ für $R > 0$ löst. [„Integralgleichungsmethode von Levi.“ Vgl. Lichtenstein, Diff.gl., S. 1295–1296 nebst Fußnote 38. Betreffe hinreichend kleiner b und $(\Delta + T)$ anstatt E vgl. Courant, Bd. 2, S. 279–282.]

Diese reduzierte Grundlösung ist nicht die Hilbertsche Parametrix, und das daran anknüpfende Verfahren nicht die Hilbertsche Parametrixmethode (wie Courant behauptet). Denn obiges Verfahren fußt auf der in (11. 2) ausgedrückten Vertauschung von Differentiation und Integration und löst erst die Differentialgleichung ohne Randbedingung. Dagegen fußt Hilberts P -Methode auf der in (11. 1) ausgedrückten partiellen Integration und löst bei sinngemäßer Verallgemeinerung auf eigentliche Randwertprobleme mit dem auf dem Rand verschwindenden P , das keiner Differentialgleichung genügt, unmittelbar das Randwertproblem. Ferner gewinnt man durch die

auf (11. 1) sich gründende *P*-Methode alle Hauptsätze, was für die von Courant angeführte Levi-Fredholmsche Methode nicht der Fall ist. (Vgl. Hilbert, Haupt, Lichtenstein Diff.gl., Hellinger, wie in den Anmerkungen zu Nr. 12a näher angegeben wird.) Courant bezeichnet die reduzierte Grundlösung auch als „Singularitätenfunktion“ (Courant, Bd. 2, S. 279). Diesen Begriff werde ich in Nr. 23a wesentlich verallgemeinern.

II. Die „Integralgleichungsmethode von Fredholm“ transformiert das Randwertproblem in eine Integralgleichung ähnlich (11. 7). Der Kern derselben wird aus der oben in I angedeuteten vollständigen Grundlösung abgeleitet. Die Integration erstreckt sich jedoch über d . Diese Methode setzt notwendigerweise voraus, daß d keine Ecken hat. (Vgl. Lichtenstein, Diff.gl., Fußnote 38 auf S. 1296, und Hell.-Toepl., S. 1346. — Betreffs Δ und hinreichend kleinem b vgl. Courant, Bd. 2, S. 269—272.)

Die Literaturergebnisse der Grundlösungsmethode besagen weniger als die in Nr. 11b erwähnten, weshalb ich hier nicht weiter darauf eingehe.

(Eingegangen am 14. I. 1941.)

Alfred Ackermann-Teubner-Gedächtnispreis.

Der von Herrn Domherrn Dr. Dr.-Ing. Alfred Ackermann-Teubner in Leipzig im Jahre 1912 der Universität Leipzig gestiftete „Alfred Ackermann-Teubner-Gedächtnispreis zur Förderung der Mathematischen Wissenschaften“ in Höhe von RM. 500.— ist für das Jahr 1941 durch das Preisgericht dem Professor an der Universität Göttingen Dr. P. ten Bruggencate für seine Arbeiten zur Erforschung der Fixsternatmosphäre zuerkannt worden.

Beweis und Verschärfung eines Satzes von Kronecker.

Von

Oskar Perron in München.

Vorbemerkung. In dieser Note handelt es sich um Polynome und algebraische Gleichungen mit Koeffizienten aus einem Körper der Charakteristik 0, so daß er den Körper der rationalen Zahlen enthält. Es könnte aber ebenso gut auch ein *unendlicher* Körper mit Primzahlcharakteristik zugelassen werden.

§ 1.

Formulierung und Beweis des Hauptsatzes.

Kronecker hat ohne ausführlichen Beweis den folgenden Satz angegeben¹⁾:

Satz 1. *Ein System von beliebig vielen Gleichungen mit $k - 1$ Unbekannten ist stets einem System von höchstens k Gleichungen äquivalent (d. h. beide Systeme haben dieselben Lösungen).*

Außer einem nicht leicht zu durchschauenden Beweis dieses Satzes von König²⁾ ist mir nur noch ein unveröffentlichter Beweis des Herrn van der Waerden bekannt³⁾. Im folgenden soll ein besonders einfacher Beweis gegeben werden, wobei noch eine merkliche Verschärfung des Satzes herauskommt. Zunächst sei das Analogon für homogene Gleichungen formuliert:

Satz 2. *Ein System von beliebig vielen homogenen Gleichungen mit k Unbekannten ist stets einem System von höchstens k homogenen Gleichungen äquivalent.*

Es ist klar, daß Satz 1 aus Satz 2 folgt, indem man eine Unbekannte gleich 1 setzt. Es genügt also, den Satz 2 zu beweisen. Nun ist das homogene Gleichungssystem

$$(1) \quad f_v(x_1, \dots, x_k) = 0 \quad (v = 1, \dots, n),$$

wenn f_v vom Grad m_v ist und $\text{Max } m_v = m$ gesetzt wird, offenbar äquivalent mit dem System

$$(2) \quad x_\alpha^{m-m_v} f_v(x_1, \dots, x_k) = 0 \quad (v = 1, \dots, n; \alpha = 1, \dots, k).$$

¹⁾ Grundzüge einer arithmetischen Theorie der algebraischen Größen, § 10. Crelles Journ. f. Mathematik, Band 92 = Werke, Band II.

²⁾ Julius König, Einleitung in die allgemeine Theorie der algebraischen Größen, S. 234ff. Leipzig 1903.

³⁾ Eine Skizze findet man im Zentralblatt für Mathematik 24, S. 276.

Hier sind die linken Seiten lauter homogene Polynome gleichen Grades. Es genügt daher, den Satz 2 für den Fall zu beweisen, daß die Gleichungen des Systems sämtlich von gleichem Grad sind. Wir beweisen gleich die folgende Verschärfung:

Satz 3. Zu n homogenen Polynomen m -ten Grades von k Variabeln

$$f_v(x_1, \dots, x_k) \quad (v = 1, \dots, n)$$

kann man stets k Linearkombinationen

$$\varphi_x(x_1, \dots, x_k) = \sum_{v=1}^n a_{xv} f_v(x_1, \dots, x_k) \quad (x = 1, \dots, k)$$

mit rationalen Koeffizienten a_{xv} so angeben, daß das Gleichungssystem

$$f_v(x_1, \dots, x_k) = 0 \quad (v = 1, \dots, n)$$

äquivalent ist dem System

$$\varphi_x(x_1, \dots, x_k) = 0 \quad (x = 1, \dots, k).$$

Beweis. Für $n \leq k$ ist nichts zu beweisen; sei also $n > k$. Es genügt, sich auf den Fall $n = k + 1$ zu beschränken. Denn, wenn man dafür den Satz hat, so kann man ihn im Fall $n > k + 1$ zunächst auf das System der ersten $k + 1$ Gleichungen anwenden, wodurch man ein äquivalentes System mit nur noch $n - 1$ Gleichungen erhält, mit dem man ebenso verfahren kann usw.

Nun besteht, wenn M eine genügend große natürliche Zahl ist, zwischen den $k + 1$ Polynomen $f_v(x_1, \dots, x_k)$ eine Abhängigkeit der Gestalt

$$(3) \quad \sum A_{\lambda_1 \lambda_2 \dots \lambda_k} f_1^{\lambda_1} f_2^{\lambda_2} \dots f_{k+1}^{\lambda_{k+1}} = 0 \quad (\lambda_1 + \lambda_2 + \dots + \lambda_{k+1} = M).$$

Denn die linke Seite ist in den f_v homogen vom Grad M , in den x_x also homogen vom Grad mM , und da ein vollständiges homogenes Polynom M -ten Grades von $k + 1$ Variabeln bzw. (mM) -ten Grades von k Variabeln $\binom{M+k}{k}$ bzw. $\binom{mM+k-1}{k-1}$ Glieder hat, müssen die $\binom{M+k}{k}$ unbekannten Koeffizienten $A_{\lambda_1 \lambda_2 \dots \lambda_k}$ lediglich einem System von $\binom{mM+k-1}{k-1}$ linearen homogenen Bedingungsgleichungen genügen, damit (3) identisch in den x_x erfüllt ist. Für genügend großes M ist aber

$$\binom{M+k}{k} > \binom{mM+k-1}{k-1},$$

so daß die Bedingungsgleichungen sich erfüllen lassen, ohne daß alle $A_{\lambda_1 \lambda_2 \dots \lambda_k}$ verschwinden.

Nun betrachten wir $k + 1$ Linearkombinationen

$$(4) \quad \varphi_x = a_{x1} f_1 + a_{x2} f_2 + \dots + a_{x, k+1} f_{k+1} \quad (x = 1, \dots, k+1)$$

mit nicht verschwindender Determinante $|a_{xv}|$, so daß sich durch Auflösung ergibt

$$(5) \quad f_v = b_{v1} \varphi_1 + b_{v2} \varphi_2 + \dots + b_{v, k+1} \varphi_{k+1} \quad (v = 1, \dots, k+1),$$

wobei auch $|b_{vx}| \neq 0$ ist. Das System $f_v = 0$ ist dann äquivalent dem System $\varphi_x = 0$. Zwischen den φ_x ergibt sich aber, indem man die Ausdrücke (5) in (3) einsetzt, eine Identität der Form

$$(6) \quad \sum B_{\lambda_1 \lambda_2 \dots \lambda_k} b_{1, k+1}^{\lambda_1} b_{2, k+1}^{\lambda_2} \dots b_{k+1, k+1}^{\lambda_k} = 0 \quad (\lambda_1 + \lambda_2 + \dots + \lambda_{k+1} = M),$$

wobei insbesondere

$$B_{00\dots 0} = \sum A_{\lambda_1 \lambda_2 \dots \lambda_k} b_{1, k+1}^{\lambda_1} b_{2, k+1}^{\lambda_2} \dots b_{k+1, k+1}^{\lambda_k}.$$

Nun kann man bekanntlich die b_{vx} als rationale Zahlen so wählen, daß die Determinante $|b_{vx}|$ und das Polynom $B_{00\dots 0}$ von 0 verschiedene Werte haben⁴⁾. Dann darf man $B_{00\dots 0} = 1$ annehmen, und die Identität (6) hat das Aussehen

$$\varphi_{k+1}^M + \varphi_{k+1}^{M-1} \sum_{v=1}^k B_v \varphi_v + \varphi_{k+1}^{M-2} \sum_{v=1}^k \sum_{\mu=1}^k B_{v\mu} \varphi_v \varphi_\mu + \dots = 0.$$

Hiernach ist klar, daß die Gleichungen

$$(7) \quad \varphi_1 = 0, \varphi_2 = 0, \dots, \varphi_k = 0$$

die Gleichung $\varphi_{k+1} = 0$ nach sich ziehen. Das System

$$\varphi_1 = 0, \varphi_2 = 0, \dots, \varphi_k = 0, \varphi_{k+1} = 0$$

und folglich auch das System

$$f_1 = 0, f_2 = 0, \dots, f_k = 0, f_{k+1} = 0$$

ist also äquivalent dem System (7). Damit ist Satz 3 bewiesen.

Der Vollständigkeit halber sei auch noch das Analogon für nicht notwendig homogene Gleichungen angegeben:

Satz 4. Zu n Polynomen von $k - 1$ Variablen

$$f_v(x_1, \dots, x_{k-1}) \quad (v = 1, \dots, n)$$

kann man stets k Linearkombinationen

$$\varphi_x(x_1, \dots, x_{k-1}) = \sum_{v=1}^n a_{xv} f_v(x_1, \dots, x_{k-1}) \quad (x = 1, \dots, k)$$

⁴⁾ Satz 29 in Band 1 meiner „Algebra“ (1927, 2. Aufl., 1932).

mit rationalen Koeffizienten a_x , so angeben, daß das Gleichungssystem

$$f_v(x_1, \dots, x_{k-1}) = 0 \quad (v = 1, \dots, n)$$

äquivalent ist dem System

$$\varphi_x(x_1, \dots, x_{k-1}) = 0 \quad (x = 1, \dots, k).$$

Zum Beweis braucht man nur $\frac{x_v}{x_k}$ statt x_v zu setzen und dann die f_v durch Multiplikation mit geeigneten Potenzen von x_k zu homogenen Polynomen gleichen Grades zu machen. Auf diese wendet man den Satz 3 an und setzt hinterher $x_k = 1$.

§ 2.

Was der Satz von Kronecker besagt und was er nicht besagt.

Wenn die obigen Sätze 1 und 2 aussagen, daß man stets mit k Gleichungen auskommt, so schließt das natürlich nicht aus, daß es in Spezialfällen schon mit weniger als k Gleichungen geht. Möglicherweise gilt sogar der Satz, daß in allen Fällen, in denen das Gleichungssystem überhaupt eine Lösung hat (im homogenen Fall eine nichttriviale Lösung), schon $k - 1$ Gleichungen genügen. Doch kann ich das nicht beweisen, während mir andererseits aber auch kein Beispiel bekannt ist, bei dem $k - 1$ Gleichungen wirklich nicht genügen. Als ich kürzlich im Fall $k = 4$ bei einem Beispiel, bei dem man 50 Jahre lang geglaubt hat, daß 4 Gleichungen nötig sind, zeigte, daß doch schon 3 Gleichungen genügen⁵⁾, begegneten meine Ausführungen bei verschiedenen Lesern einem merkwürdigen Mißverständnis. Man sagte, meine Widerlegung sei nur dadurch zustande gekommen, daß ich durch Nichtberücksichtigung der Vielfachheit einer Lösung von der „üblichen“ Auffassung abgewichen sei. Dem scheint die Meinung zugrunde zu liegen, daß der Kroneckersche Satz der „üblichen“ Auffassung entspricht und erst ich mich einer Abweichung davon schuldig gemacht habe. In Wahrheit nimmt aber auch Kronecker keine Rücksicht auf Vielfachheit, was ich damals vorsichtshalber sogar ausdrücklich sagte, und es hat sich ja um ein Beispiel zum Satz von Kronecker gehandelt. Wer von diesem Satz spricht, muß, wenn er nicht ausdrücklich das Gegenteil sagt, notwendig den gleichen Standpunkt wie Kronecker einnehmen. Das taten die von mir zitierten Autoren und tat natürlich auch ich.

Die Vielfachheit *kann* im Kroneckerschen Satz schon deshalb gar nicht berücksichtigt sein, weil es eine Definition, die diesen Begriff in einer für alle Fälle gültigen Weise eindeutig festlegt, zu Kroneckers Zeit nicht gab und

⁵⁾ Über das Vahlensche Beispiel zu einem Satz von Kronecker, Math. Zeitschr. 47, S. 318–324.

übrigens auch heute nicht gibt. Man könnte aber vielleicht an folgendes denken:

Jedes Polynomsystem f_1, \dots, f_n ist Basis eines Polynomideals $J = (f_1, \dots, f_n)$, speziell eines H -Ideals im Fall homogener Polynome. Die Lösungen des Gleichungssystems $f_i = 0$ bilden die Nullstellenmannigfaltigkeit (kurz NM) des Ideals J . Es gibt unendlich viele Ideale mit der gleichen NM ; z. B. haben die Ideale

$$(8) \quad (f_1, f_2), (f_1^r, f_2^s), (f_1^r, f_2^s, f_1^p f_2^q) \quad (r < p; s < q)$$

und noch viele andere Typen dieselbe NM . Unter diesen unendlich vielen Idealen ist das zur NM „zugehörige Ideal“ besonders ausgezeichnet; es ist die Gesamtheit *aller* Polynome, die in jedem Punkt der NM verschwinden. Jedes dieser unendlich vielen verschiedenen Ideale repräsentiert die NM in einer irgendwie anderen individuellen Weise. Eine erschöpfende Beschreibung der Unterschiede ist mir nicht bekannt; man wird aber den Hauptunterschied vielleicht in einer Art Vielfachheit erblicken; z. B. wird man geneigt sein, zu sagen, von den drei Idealen (8) geben die letzten beiden die Punkte der NM in höherer Vielfachheit als das erste (aber in welcher Vielfachheit gibt es das dritte?). Und von dem *zugehörigen* Ideal wird man vielleicht sagen, daß es die Punkte der NM in der „richtigen“ (= minimalen) Vielfachheit gibt. Man kann nun, wenn man die Forderung der Erhaltung einer nicht klar definierten Vielfachheit stellt, etwa gleich an die Erhaltung des Ideals denken, weil man bei Nichterhaltung ja doch irgendwie „abweicht“. Das wäre wenigstens etwas klar Definiertes, aber der Kroneckersche Satz wäre dann ganz falsch. Der Satz sagt nämlich folgendes aus:

Unter den unendlich vielen Idealen mit derselben NM gibt es eines, das eine Basis von höchstens k Polynomen hat.

Er sagt aber nicht, daß *jedes* Ideal eine Basis von höchstens k Polynomen hat, auch nicht einmal, daß das zu einer NM *zugehörige* Ideal eine Basis von höchstens k Polynomen hat. Das wäre ganz falsch. Z. B. bilden die $\binom{k-1}{2}$ Determinanten der Matrix

$$\begin{vmatrix} x_1 & x_2 & \dots & x_{k-1} \\ x_2 & x_3 & \dots & x_k \end{vmatrix}$$

ein H -Ideal, dessen NM die folgende ist:

$$x_1 = \varrho^{k-1}, x_2 = \varrho^{k-2} \sigma, \dots, x_k = \sigma^{k-1}.$$

Man sieht leicht, daß das H -Ideal das *zugehörige* Ideal ist und daß es keine Basis von weniger als $\binom{k-1}{2}$ Polynomen besitzt; denn die $\binom{k-1}{2}$ Determinanten sind linear unabhängig. Die NM aber kann man nach dem Kronecker-

sehen Satz schon durch k , ja sogar, wie ich an anderer Stelle gezeigt habe⁶⁾, schon durch $k - 1$ homogene Gleichungen darstellen. Die betreffenden H -Ideale sind natürlich andere, aber die NM ist dieselbe⁷⁾.

§ 3.

Über die Abhängigkeit homogener Polynome gleichen Grades.

Es soll bei dieser Gelegenheit noch die obige Formel (3) verschärft werden zu einem Satz, dessen Nützlichkeit sich an anderer Stelle zeigen wird.

Satz 5. Zwischen $k + 1$ homogenen Polynomen m -ten Grades

$$f_1(x_1, \dots, x_k), \dots, f_{k+1}(x_1, \dots, x_k)$$

besteht eine Abhängigkeit

$$\sum A_{\lambda_1 \lambda_2 \dots \lambda_k} f_1^{\lambda_1} f_2^{\lambda_2} \dots f_{k+1}^{\lambda_{k+1}} = 0,$$

homogen vom Grad m^{k-1} (also $\lambda_1 + \lambda_2 + \dots + \lambda_{k+1} = m^{k-1}$).

Die bei Formel (3) angewandte Beweismethode versagt hier, weil die Zahl der Bedingungsgleichungen, von wenigen Ausnahmefällen abgesehen, jetzt größer ist als die der unbekannten Koeffizienten. Zum Beweis von Satz 5 benötigen wir den

Hilfssatz. Die Koeffizienten der k Polynome von k Variablen

$$f_v(x_1, \dots, x_k) = a_v x_1^{m_v} + \dots \quad (v = 1, \dots, k)$$

von den positiven Graden m_v seien, soweit sie nicht verschwinden, Unbestimmte und speziell die k angeschriebenen Koeffizienten a_v mögen nicht verschwinden. Dann läßt sich jedes Polynom $F(x_1, \dots, x_k)$ höchstens n -ten Grades auf eine und nur eine Art in die Form setzen:

$$F(x_1, \dots, x_k) = \sum C_{q_1 \dots q_k}^{r_1 \dots r_k} f_1^{q_1} \dots f_k^{q_k} x_1^{r_1} \dots x_k^{r_k},$$

wenn dabei verlangt wird, daß in jedem Glied

$$m_1 q_1 + \dots + m_k q_k + r_1 + \dots + r_k \leq n,$$

$$0 \leq r_1 \leq m_1 - 1, \dots, 0 \leq r_k \leq m_k - 1$$

sein soll. Die Koeffizienten $C_{q_1 \dots q_k}^{r_1 \dots r_k}$ gehören dem durch die Koeffizienten der Polynome f_1, \dots, f_k, F bestimmten Körper an.

⁶⁾ Über die Bedingungen, daß eine binäre Form n -ten Grades eine n -te Potenz ist, und über die rationale Kurve n -ter Ordnung im E_n . Math. Annalen 118, S. 305–309.

⁷⁾ Es gibt sogar Ideale, bei denen die Mindestzahl der Basispolynome beliebig groß ist (bei fester Variablenzahl). Ein Beispiel findet man bei F. S. Macaulay, The algebraic theory of modular systems, art. 34 (Cambridge 1916).

Der Beweis des Hilfssatzes ist wörtlich derselbe wie bei Satz 58 in Band I meiner „Algebra“ (1927, 2. Aufl. 1932). Dort wurde zwar bei Formulierung des Satzes von *allen* Koeffizienten der f , verlangt, daß sie nicht verschwinden; beim Beweis wurde aber nur das Nichtverschwinden der oben angeschriebenen benutzt.

Wir wenden den Hilfssatz an auf k vollständige homogene Polynome m -ten Grades f_1, \dots, f_k mit Unbestimmten als Koeffizienten und auf das Polynom

$$F(x_1, \dots, x_k) = f_{k+1}(x_1, \dots, x_k) x_1^{s_1} \dots x_k^{s_k},$$

wobei auch f_{k+1} ein homogenes Polynom vom Grad m ist und die s , irgendein den Forderungen

$$(9) \quad 0 \leq s_r \leq m-1, \quad \sum_{r=1}^k s_r \equiv 0 \pmod{m}$$

genügendes System ganzer Zahlen sind. Der Hilfssatz liefert dann eine Formel der Gestalt

$$(10) \quad f_{k+1} \cdot x_1^{s_1} \dots x_k^{s_k} = \sum_{r_1, \dots, r_k} \Phi_{r_1, \dots, r_k}^{s_1, \dots, s_k}(f_1, \dots, f_k) \cdot x_1^{r_1} \dots x_k^{r_k},$$

$$0 \leq r_v \leq m-1, \quad \sum_{v=1}^k r_v \equiv 0 \pmod{m}.$$

Denn da die linke Seite einen durch m teilbaren Grad hat, können rechts nur Glieder kommen, in denen $\sum r_i$ durch m teilbar ist. Zugleich sieht man, daß $\phi_{r_1, \dots, r_k}^{s_1, \dots, s_k}$ ein Polynom von f_1, \dots, f_k ist vom Grad

$$\frac{1}{m} (m + \sum s_i - \sum r_i).$$

Die Anzahl der Systeme s_1, \dots, s_k , die den Bedingungen (9) genügen, ist gleich m^{k-1} . Setzt man $m^{k-1} = l$ und bezeichnet die l Potenzprodukte $x_1^{s_1} \dots x_k^{s_k}$ in irgendeiner Reihenfolge mit U_1, \dots, U_l , so läßt sich die Formel (10) folgendermaßen schreiben:

$$(11) \quad f_{k+1} \cdot U_k = \sum_{\mu=1}^l \Phi_{k\mu}(f_1, \dots, f_k) \cdot U_\mu \quad (k = 1, \dots, l).$$

Hieraus folgt

$$(12) \quad \begin{vmatrix} \Phi_{11} - l_{k+1} & \dots & \Phi_{1l} \\ \dots & \dots & \dots \\ \Phi_{l1} & \dots & \Phi_{ll} - l_{k+1} \end{vmatrix} = 0.$$

Bezeichnet man mit mg_i den Grad von U_i , so ist $\Phi_{i,\mu}$ nach (11) ein Polynom von f_1, \dots, f_k , homogen vom Grad $1 + g_i - g_\mu$ oder identisch 0. In der ausgerechneten Determinante (12) ist daher jedes Glied ein Polynom von

f_1, \dots, f_k, f_{k+1} , homogen vom Grad l oder identisch 0. Die Formel (12) stellt somit eine Abhängigkeit der $k+1$ Polynome f_i dar von der in Satz 5 behaupteten Gestalt. Damit ist Satz 5 zunächst für den Fall bewiesen, daß die Koeffizienten von f_1, \dots, f_k Unbestimmte sind, also erst recht für den Fall, daß die Koeffizienten aller Polynome f_1, \dots, f_k, f_{k+1} Unbestimmte sind.

Nun involviert die Existenz einer Identität von der in Satz 5 behaupteten Gestalt für die unbekannten Koeffizienten $A_{i_1 i_2 \dots i_k}$ ein System linearer homogener Bedingungsgleichungen, und zwar mehr Gleichungen als Unbekannte. Die Koeffizienten dieses Gleichungssystems sind Polynome der Koeffizienten der f_i . Sind diese nun Unbestimmte, so hat das Gleichungssystem eine nichttriviale Lösung, da wir ja für diesen Fall die Existenz der in Satz 5 behaupteten Identität bewiesen haben. Der Rang der Matrix des Gleichungssystems ist daher kleiner als die Anzahl p der Unbekannten, so daß alle p -reihigen Determinanten verschwinden. Ersetzt man jetzt die Koeffizienten, die seither Unbestimmte waren, durch irgendwelche speziellen Größen, so müssen diese Determinanten erst recht verschwinden. Das Gleichungssystem hat also immer noch eine nichttriviale Lösung für die Koeffizienten $A_{i_1 i_2 \dots i_k}$, und mit diesen gilt dann die in Satz 5 behauptete Identität.

(Eingegangen am 5. I. 1942.)

Induktion partiell stetiger Funktionen.

Von

Jos. Novák in Brünn.

Jedes System \mathfrak{S} reeller Funktionen, die in einer abstrakten Menge M definiert sind, induziert in M eine Topologie, die sogenannte Induktion des Systems \mathfrak{S} . Diese Induktion erfüllt im allgemeinen nur das erste und das dritte Hausdorffsche Umgebungsaxiom (A) und (C), wogegen das zweite (B) dann und nur dann erfüllt ist, falls die Summe zweier in M stetiger Funktionen eine in M stetige Funktion ist. In der vorliegenden Arbeit werden wir uns mit der Induktion w partiell stetiger Funktionen von zwei reellen Veränderlichen befassen, d. i. derartiger Funktionen $f(x, y)$, daß die Funktionen einer reellen Veränderlichen $f(x_0, y)$ und $f(x, y_0)$ für jeden Punkt (x_0, y_0) stetig sind. Verstehen wir unter der „Konvergenz“ einer Punktfolge gegen einen Punkt die Eigenschaft, daß jede Umgebung dieses Punktes fast alle Punkte der Folge enthält, so „konvergiert“ die Punktfolge $\{(x_n, y_n)\}_{n=1}^{\infty}$ gegen den Punkt (x_0, y_0) in der Induktion w dann und nur dann, wenn fast alle Punkte (x_n, y_n) auf der waagerechten oder lotrechten Geraden oder auf diesen beiden Geraden gegen den Punkt (x_0, y_0) im gewöhnlichen Sinne konvergieren. Es möchte also scheinen, daß die Induktion w durch diese Eigenschaft der „Konvergenz“ ganz beschreibbar ist. Diese „Konvergenz“ bestimmt jedoch die sogenannte L -Induktion u partiell stetiger Funktionen, die von ganz verschiedener Natur ist. Auch die stärkere C -Modifikation der L -Induktion u , die alle vier Hausdorffschen Umgebungsaxiome erfüllt, ist keineswegs mit der Induktion w topologisch äquivalent, denn diese ist regulär, jene jedoch nicht [Satz 1]. Die Induktion w ist in keinem Punkte normal [Satz 3] und deshalb auch nicht mit der Euklidischen Topologie homöomorph.

Die Arbeit besteht aus zwei Teilen. Der erste Teil handelt über die L -Induktion u partiell stetiger Funktionen, die als Kreuztopologie bezeichnet wird, und über die C -Modifikation dieser Topologie. In dem zweiten Teile wird die Induktion w partiell stetiger Funktionen untersucht. Es werden die Beziehungen zwischen diesen beiden Topologien sowie auch ihre Stellen in der Klassifikation der topologischen Räume bestimmt.

1.

Im topologischen Seminar vom 14. November 1939 legte E. Čech die Frage vor: *Es sei u diejenige Topologie¹⁾ in der Ebene R_2 , in der die Umgebungen die Gestalt eines Kreuzes haben. Für welche Ordinalzahl ξ gibt es eine derartige Menge $M \subset R_2$, daß $u^{\xi+1}M = u^\xi M \neq u^\zeta M$ für alle $\zeta < \xi$? Die Antwort lautet: Zu jeder Ordinalzahl $\xi \leq \omega_1$ (ω_1 ist die erste un abzählbare Ordinalzahl) gibt es in der Ebene eine Menge mit den verlangten Eigenschaften.*

Der Beweis dieser Behauptung ist in den zwei Konstruktionen I und II enthalten.

Konstruktion I. *Es sei $(a, b) \in R_2$. Zu jeder abzählbaren Ordinalzahl α und zu jedem gegebenen abzählbaren System \mathfrak{I} waagerechter und lotrechter Geraden²⁾, von denen keine den Punkt (a, b) enthält, gibt es eine abzählbare Punktmenge N mit folgenden Eigenschaften:*

$$[1\alpha] \quad (a, b) \in u^{\alpha+1}N - u^\alpha N.$$

$$[2\alpha] \quad u^{\alpha+1}N \text{ ist eine abzählbare und abgeschlossene Menge.}$$

$$[3\alpha] \quad \text{Kein Punkt der Menge } u^{\alpha+1}N \text{ liegt auf irgendeiner Geraden des Systems } \mathfrak{I}.$$

Es ist möglich, die Konstruktion so durchzuführen, daß die Menge N innerhalb eines Kreises von beliebig vorher gegebenem Radius ε liegt, der in dem Punkte (a, b) die waagerechte (oder lotrechte)-Gerade berührt.

Wir beweisen zunächst den **Hilfssatz**: *Es sei $(c, d) \in R_2$; die wachsende Zahlenfolge $\{d_m\}$ konvergiere gegen die Zahl d . Wir zeichnen um jeden Punkt $(c + r_m, d_m) \in R_2$ in der Ebene einen Kreis vom Radius $r_m = \frac{1}{2} \min \{d_{m+1} - d_m, d_m - d_{m-1}\}$, $m = 1, 2, \dots$ der die lotrechte Gerade im Punkte (c, d_m) berührt.*

¹⁾ Unter einem topologischen Raume verstehen wir eine solche Menge, in der jedem Punkte ein definierendes Umgebungssystem zugeordnet ist, wobei das erste Hausdorffsche Umgebungsaxiom (A) erfüllt ist, anders gesagt eine solche Menge, in der jeder Teilmenge M die sogenannte „abgeschlossene“ Hülle \bar{M} derart zugeordnet ist, daß folgendes Hülleaxiom gilt: $\bar{\bar{a}} = \bar{a}$; $M \subset \bar{M}$: ist $M_1 \subset M_2$, so ist $\bar{M}_1 \subset \bar{M}_2$. Diese Zuordnungsregel nennen wir Topologie und bezeichnen sie mit den Buchstaben u, v, w, \dots , usw. Wir setzen $u^0 M = M$, $u^1 M = u M = \bar{M}$; ist ξ eine Ordinalzahl, so bedeutet das Symbol u^ξ die Topologie, die der Teilmenge M die Hülle $u^\xi M = u(u^{\xi-1} M)$, falls ξ eine isolierte Zahl, und $u^\xi M = \sum_{\zeta < \xi} u^\zeta M$, falls ξ eine Grenzzahl $\neq 0$ ist, zuordnet.

²⁾ Unter einer waagerechten (lotrechten) Geraden in der Ebene verstehen wir jede Parallele zur Achse x (y).

Diesen Kreis samt dem Innern bezeichnen wir mit K_m . Wir wählen eine Menge $M_m \subset K_m$. Ist γ eine beliebige Ordinalzahl, so ist entweder

$$u^\gamma \sum_{m=1}^{\infty} M_m = \sum_{m=1}^{\infty} u^\gamma M_m$$

oder

$$u^\gamma \sum_{m=1}^{\infty} M_m = \sum_{m=1}^{\infty} u^\gamma M_m + (c, d).$$

Beweis. Da $M_m \subset \sum_{m=1}^{\infty} M_m$, so gilt nach dem Hülleaxiom: $u^\gamma M_m \subset u^\gamma \sum_{m=1}^{\infty} M_m$ für jedes $m = 1, 2, \dots$, so daß $\sum_{m=1}^{\infty} u^\gamma M_m \subset u^\gamma \sum_{m=1}^{\infty} M_m$; es bleibt also noch zu zeigen, daß $u^\gamma \sum_{m=1}^{\infty} M_m \subset \sum_{m=1}^{\infty} u^\gamma M_m + (c, d)$. Diese Beziehung beweisen wir durch die transfinite Induktion. Die Beziehung gilt für $\gamma = 0$. Wir setzen ihre Geltung für alle Ordinalzahlen, die $< \gamma$ sind, voraus, und beweisen sie auch für γ . Wir unterscheiden zwei Fälle.

1. $\gamma > 0$ ist eine Grenzzahl. Nach der Definition der Topologie u^γ , nach Voraussetzung und da für Mengensummen das kommutative Gesetz gilt, ist

$$\begin{aligned} u^\gamma \sum_{m=1}^{\infty} M_m &= \sum_{\lambda < \gamma} u^\lambda \sum_{m=1}^{\infty} M_m \subset \sum_{\lambda < \gamma} \left[\sum_{m=1}^{\infty} u^\lambda M_m + (c, d) \right] = \\ &= \sum_{m=1}^{\infty} \sum_{\lambda < \gamma} u^\lambda M_m + (c, d) = \sum_{m=1}^{\infty} u^\gamma M_m + (c, d). \end{aligned}$$

2. γ ist eine isolierte Zahl. In diesem Falle gilt:

$$\begin{aligned} u^\gamma \sum_{m=1}^{\infty} M_m &= u \left[u^{\gamma-1} \sum_{m=1}^{\infty} M_m \right] \subset u \left[\sum_{m=1}^{\infty} u^{\gamma-1} M_m + (c, d) \right] = \\ &= u \sum_{m=1}^{\infty} u^{\gamma-1} M_m + (c, d). \end{aligned}$$

Es bleibt deshalb noch zu beweisen, daß

$$u \sum_{m=1}^{\infty} u^{\gamma-1} M_m \subset \sum_{m=1}^{\infty} u^\gamma M_m + (c, d).$$

Das ist sicher der Fall. Ist nämlich $(x, y) \in u \sum_{m=1}^{\infty} u^{\gamma-1} M_m$, so hat jede Kreuzumgebung gemeinsame Punkte mit der Menge $\sum_{m=1}^{\infty} u^{\gamma-1} M_m$. Entweder gibt es nun einen Index p derart, daß in jeder Kreuzumgebung des Punktes (x, y) Punkte der Menge $u^{\gamma-1} M_p$ liegen; dann ist also $(x, y) \in u^\gamma M_p$ und daher auch $(x, y) \in \sum_{m=1}^{\infty} u^\gamma M_m$. Oder ein solcher Index existiert nicht; das ist nur so möglich, daß $(x, y) = (c, d)$. In der Tat, da die Menge K_m in der Eukli-

dischen Topologie abgeschlossen ist und da die Euklidische Topologie stärker ³⁾ als jede Topologie u^ξ (ξ ist eine beliebige Ordinalzahl) ist, so gilt $u^{\gamma-1} M_m \subset K_m$ für $m = 1, 2, \dots$

Die Konstruktion I wird mittels transfiniter Induktion durchgeführt werden. Wir setzen zu diesem Zwecke voraus, daß wir zu einem gegebenen Geradensystem \mathfrak{T} und zu jeder Ordinalzahl, die kleiner als α ist, eine abzählbare Menge von den angegebenen Eigenschaften schon konstruiert haben und wir konstruieren eine solche Menge für die Zahl α .

Es sei $\varepsilon > 0$ gegeben. Es sei $\{b_m\}$ eine wachsende Folge reeller Zahlen, die gegen den Grenzwert b konvergiert, wobei $b - \varepsilon < b_1$; wir setzen weiter voraus, daß kein Punkt (a, b_m) auf einer Geraden des Systems \mathfrak{T} liegt. Um jeden der Punkte $(a + r_m, b_m)$ zeichnen wir einen Kreis mit dem Radius $r_m = \frac{1}{2} \min \{b_{m+1} - b_m, b_m - b_{m-1}\}$, $m = 1, 2, \dots$, wo $b_0 = b - \varepsilon$, so daß der Kreis die lotrechte Gerade im Punkte (a, b_m) berührt. Die Menge der Punkte auf diesem Kreise und in dessen Innerem bezeichnen wir mit K_m .

Ist $\alpha = 0$, so hat die Menge $N = \sum_{m=1}^{\infty} (a, b_m)$ alle drei Eigenschaften $[1\alpha]$, $[2\alpha]$ und $[3\alpha]$. Es sei nun $\{\alpha_m\}$ eine Folge von Ordinalzahlen, so daß $\lim (\alpha_m + 1) = \alpha$, falls α eine Grenzzahl > 0 und $\alpha_m = \alpha - 1$, wenn α eine isolierte Zahl⁴⁾ ist. Nach der Induktionsvoraussetzung gibt es zu dem Punkte (a, b_m) und zu dem Geradensystem \mathfrak{T} im Innern des Kreises K_m eine abzählbare Menge $N_m \subset R_{\frac{1}{2}}$, die alle drei Eigenschaften $[1\alpha_m]$, $[2\alpha_m]$ und $[3\alpha_m]$ hat.

Wir setzen $N = \sum_{m=1}^{\infty} N_m$.

1. Es ist zuerst $(a, b) \in u^{\alpha+1} N$. Tatsächlich, jede Kreuzumgebung des Punktes (a, b) hat gemeinsame Punkte mit der Menge $\sum_{m=1}^{\infty} (a, b_m)$, so daß $(a, b) \in u \sum_{m=1}^{\infty} (a, b_m)$. Da $\sum_{m=1}^{\infty} (a, b_m) \subset \sum_{m=1}^{\infty} u^{\alpha_m+1} N = u^{\alpha} N$, so gilt:

$$u \sum_{m=1}^{\infty} (a, b_m) \subset u^{\alpha+1} N.$$

Weiter ist (a, b) non $\in u^{\alpha} N$. Dies beweisen wir wieder durch transfinite Induktion. Da (a, b) non $\in u^0 N = N$, setzen wir $\alpha > 0$ voraus. Es sei $\beta < \alpha$. Es sei (a, b) non $\in u^{\beta} N$ für jede Ordinalzahl $\zeta < \beta$. Ist β eine Grenzzahl, dann ist (a, b) non $\in \sum_{\zeta < \beta} u^{\zeta} N$, d. i. (a, b) non $\in u^{\beta} N$. Wir setzen nun

³⁾ Die Topologie u_1 ist schwächer als die Topologie u_2 oder die Topologie u_2 ist stärker als die Topologie u_1 , falls jede Umgebung in der Topologie u_2 eine Umgebung desselben Punktes in der Topologie u_1 enthält. Es gilt der Satz: Die Topologie u_1 ist schwächer als die Topologie u_2 dann und nur dann, wenn $u_1 M \subset u_2 M$ für jede Unter-
menge M .

⁴⁾ In diesem zweiten Falle bedeutet das Symbol $\lim (\alpha_m + 1)$ die Zahl α .

voraus, daß β eine isolierte Zahl ist. Den Beweis setzen wir indirekt fort. Wir machen die Voraussetzung $(a, b) \in u^\beta N$. Da nach der Induktionsvoraussetzung $(a, b) \notin u^{\beta-1} N$ und da die Kreuztopologie eine L -Topologie ist⁵⁾, gibt es in der Menge $u^{\beta-1} N$ eine Punktfolge, die gegen den Punkt (a, b) konvergiert; das ist nur dadurch möglich, daß die Punktfolge eine Teilfolge der Folge $\{(a, b_m)\}$ ist. Es sei $(a, b_p) \in u^{\beta-1} N$ ein Punkt, der zu dieser Punktfolge gehört, wobei $\beta - 1 < \alpha_p$ gelte. Es ist $u^{\beta-1} N_p \subset u^{\alpha_p} N_p$ und $u^{\beta-1} N_m \subset K_m$. Der Punkt (a, b_p) liegt also in keiner Menge $u^{\beta-1} N_m$, daher auch nicht in der Menge $\sum_{m=1}^{\infty} u^{\beta-1} N_m$. Nach dem Hilfssatz ist $u^{\beta-1} N = \sum_{m=1}^{\infty} u^{\beta-1} N_m$. Daher ist $(a, b_p) \notin u^{\beta-1} N$, was einen Widerspruch ergibt.

2. Die Menge $u^{\alpha+1} N$ ist abzählbar; denn jede Menge $u^{\alpha+1} N_m = u^{\alpha_m+1} N_m$ ist nach der Voraussetzung abzählbar und nach dem Hilfssatz ist $u^{\alpha+1} N = \sum_{m=1}^{\infty} u^{\alpha+1} N_m + (a, b)$. Die Menge $u^{\alpha+1} N$ ist abgeschlossen. Es ist nämlich nach demselben Hilfssatz:

$$\begin{aligned} u^{\alpha+2} N &= u[u^{\alpha+1} N] = u\left[\sum_{m=1}^{\infty} u^{\alpha+1} N_m + (a, b)\right] = u\left[\sum_{m=1}^{\infty} u^{\alpha} N_m + (a, b)\right] = \\ &= u[u^{\alpha} N + (a, b)] = u^{\alpha+1} N + (a, b) = u^{\alpha+1} N. \end{aligned}$$

3. Weil $u^{\alpha+1} N = \sum_{m=1}^{\infty} u^{\alpha+1} N_m + (a, b) = \sum_{m=1}^{\infty} u^{\alpha_m+1} N_m + (a, b)$ und da jede Gerade des Systems \mathfrak{T} weder einen Punkt der Menge $u^{\alpha_m+1} N_m$ noch den Punkt (a, b) enthält, hat auch die Menge $u^{\alpha+1} N$ dieselbe Eigenschaft.

Die Menge N liegt im Innern des Kreises vom Radius ε ; dieser Kreis berührt die waagerechte Gerade im Punkte (a, b) . Eine Punktmenge mit denselben Eigenschaften [1 α], [2 α] und [3 α] kann man so konstruieren, daß sie innerhalb eines Kreises vom Radius ε liegt, der die lotrechte Gerade im Punkte (a, b) berührt.

Konstruktion II. In der Ebene gibt es eine derartige Menge $Q \subset R_2$, daß $u^\xi Q \neq u^{\xi+1} Q$ für jede abzählbare Ordinalzahl ξ und $u^\beta Q = u^{\beta+1} Q$ für jede unabzählbare Ordinalzahl β .

Zum Beweis benutzen wir die Methode der transfiniten Konstruktion. Es sei α eine beliebige abzählbare Ordinalzahl und für jedes $\zeta < \alpha$ sei nach

⁵⁾ Das folgt aus dem Satz: Jeder topologische Raum, der den Hausdorffschen Umgebungsaxiomen (B) und (D) und auch dem ersten Abzählbarkeitsaxiom genügt, ist ein L -Raum. Vgl. F. Hausdorff, Grundzüge der Mengenlehre, S. 263. Leipzig 1914.

der Konstruktion I ein Punkt $(a_\zeta, b_\zeta) \in R_2$ und die Menge $Q(\zeta) \subset R_2$ so gewählt, daß

$$[1 \zeta] \quad (a_\zeta, b_\zeta) \in u^{\zeta+1}Q(\zeta) - u^\zeta Q(\zeta).$$

$$[2 \zeta] \quad u^{\zeta+1}Q(\zeta) \text{ ist eine abzählbare und abgeschlossene Menge.}$$

$$[3^* \zeta] \quad u^{\zeta+1}Q(\zeta) \cdot u^{\lambda+1}Q(\lambda) = \emptyset \text{ für jedes } \zeta < \alpha, \lambda < \alpha, \zeta \neq \lambda.$$

Wir legen durch jeden Punkt der abzählbaren Menge $\sum_{\zeta < \alpha} u^{\zeta+1}Q(\zeta)$ eine waagerechte und eine lotrechte Gerade. Da das System dieser Geraden abzählbar ist, gibt es in der Ebene einen Punkt (a_α, b_α) , der auf keiner Geraden dieses Systems liegt. Nach der Konstruktion I gibt es zur Ordinalzahl α , zum Punkte (a_α, b_α) und zu diesem Geradensystem eine Menge $Q(\alpha)$, welche die Eigenschaften $[1 \alpha]$, $[2 \alpha]$ und $[3 \alpha]$ hat, so daß die Eigenschaften $[1 \zeta]$, $[2 \zeta]$ und $[3^* \zeta]$ auch für $\zeta = \alpha$ erfüllt sind. Wir setzen $Q = \sum_{\zeta < \omega_1} Q(\zeta)$. Da $Q(\zeta) \subset u^{\zeta+1}Q(\zeta)$, so ist $Q(\zeta) \cdot Q(\lambda) = \emptyset$ für jedes $\zeta < \omega_1, \lambda < \omega_1, \zeta \neq \lambda$.

Wir beweisen zunächst die Beziehung $u^\delta Q = \sum_{\zeta < \omega_1} u^\delta Q(\zeta)$ für jede Ordinalzahl δ . Zu diesem Zwecke setzen wir voraus, daß diese Beziehung für alle Ordinalzahlen $\eta < \delta$ gilt. Die Beziehung ist für $\delta = 0$ offenbar erfüllt. Ist nun $\delta > 0$ eine Grenzzahl, dann ist $u^\delta Q = \sum_{\eta > \delta} u^\eta Q = \sum_{\eta < \delta} \sum_{\zeta < \omega_1} u^\eta Q(\zeta) = \sum_{\zeta < \omega_1} (\sum_{\eta < \delta} u^\eta Q(\zeta)) = \sum_{\zeta < \omega_1} u^\delta Q(\zeta)$. Ist δ eine isolierte Zahl, so ist nach Voraussetzung $u^{\delta-1}Q = \sum_{\zeta < \omega_1} u^{\delta-1}Q(\zeta)$, so daß $u^\delta Q = u(\sum_{\zeta < \omega_1} u^{\delta-1}Q(\zeta))$. Liegt also der Punkt $(x, y) \in R_2$ in der Menge $u^\delta Q$, so hat jede seiner Kreuzumgebungen gemeinsame Punkte mit der Menge $\sum_{\zeta < \omega_1} u^{\delta-1}Q(\zeta)$; da die Kreuztopologie eine L -Topologie ist, gibt es in dieser Menge eine Punktfolge $\{(x^m, y^m)\}_{m=1}^\infty$, die gegen den Punkt (x, y) konvergiert, und zwar derart, daß jeder Punkt dieser Folge entweder auf der waagerechten oder auf der lotrechten Geraden liegt; es gibt daher eine Ordinalzahl ζ' der Eigenschaft, daß $(x^m, y^m) \in u^{\delta-1}Q(\zeta')$ für jedes $m = 1, 2, \dots$, also $(x, y) \in u^\delta Q(\zeta')$. Daher ist $u^\delta Q \subset \sum_{\zeta < \omega_1} u^\delta Q(\zeta)$. Da auch nach dem Hülleaxiom $\sum_{\zeta < \omega_1} u^\delta Q(\zeta) \subset u^\delta Q$, so ist damit die oben angeführte Beziehung bewiesen.

Es sei nun ξ eine beliebige abzählbare Ordinalzahl; wir beweisen, daß $(a_\xi, b_\xi) \in u^{\xi+1}Q - u^\xi Q$. Weil $(a_\xi, b_\xi) \in u^{\xi+1}Q(\xi)$ und da $Q(\xi) \subset Q$, so folgt aus dem Hülleaxiom, daß $(a_\xi, b_\xi) \in u^{\xi+1}Q$. Den Beweis setzen wir indirekt fort. Wir setzen voraus, daß $(a_\xi, b_\xi) \in u^\xi Q$, und folgern daraus einen Widerspruch. Nach der im vorigen Abschnitt bewiesenen Beziehung gilt $u^\xi Q = \sum_{\tau < \omega_1} u^\xi Q(\tau)$. Es gibt also eine solche Ordinalzahl τ' , daß $(a_\xi, b_\xi) \in u^\xi Q(\tau')$. Es ist $u^\xi Q(\tau') \subset u^{\tau'+1}Q(\tau')$. In der Tat, entweder ist $\xi < \tau' + 1$ und

dann gilt die Beziehung, denn $u^0 Q(\tau') \subset u^1 Q(\tau') \subset \dots \subset u^\xi Q(\tau') \subset u^{\xi+1} Q(\tau') \subset \dots \subset u^{\tau'+1} Q(\tau')$; oder es ist $\tau' + 1 \leq \xi$ und dann ist $u^\xi Q(\tau') = u^{\tau'+1} Q(\tau')$, denn die Menge $u^{\tau'+1} Q(\tau')$ ist wegen $[2 \tau']$ abgeschlossen. Es ist daher $(a_\xi, b_\xi) \in u^{\xi+1} Q(\xi) \cdot u^{\tau'+1} Q(\tau')$, so daß $\tau' = \xi$ vermöge $[3^* \xi]$. Daraus folgt $(a_\xi, b_\xi) \in u^\xi Q(\xi)$; das ist jedoch im Widerspruch zu $[1 \xi]$.

Damit ist der erste Teil, nämlich die Beziehung $u^\xi Q \neq u^{\xi+1} Q$ für jede abzählbare Ordinalzahl ξ , bewiesen.

Wir beweisen nun, daß die Menge $u^{\omega_1} Q$ abgeschlossen ist. Diese Tatsache folgt aus einem allgemeinen Satz⁶⁾ über L -Räume: Es sei S eine Menge in L -Räumen R mit der Topologie v . Dann gilt $v^{\omega_1+1} S = v^{\omega_1} S$. In der Tat, ist $z \in v^{\omega_1+1} S$, dann gibt es eine Punktfolge $\{z^m\}$, $z^m \in v^{\omega_1} S = \sum_{\xi < \omega_1} v^\xi S$, die gegen den Punkt z konvergiert, wobei $z^m \in v^{\xi_m} S$, $\xi_m < \xi_{m+1}$ und $\xi_m < \omega_1$ für jedes $m = 1, 2, \dots$. Offenbar ist $z \in v^{\lim \xi_m + 1} S \subset v^{\omega_1} S$, so daß $v^{\omega_1+1} S \subset v^{\omega_1} S$, womit auch der zweite Teil bewiesen ist.

Aus der Konstruktion II schließen wir, daß die Topologie u^ξ , wo ξ eine abzählbare Ordinalzahl ist, keine C -Modifikation⁷⁾ der Kreuztopologie ist. Erst die Topologie u^{ω_1} , die wir mit v bezeichnen wollen, ist die C -Modifikation.

Satz 1. Der Raum R_2 mit der Topologie v ist ein Hausdorffscher, jedoch kein regulärer Raum⁸⁾.

Beweis. Ist $\bar{S} \subset R_2$ und ist S eine Euklidische Hülle der Menge S , so ist $u^\gamma S \subset \bar{S}$ für jede Ordinalzahl γ . Daraus folgt, daß v schwächer ist als die Euklidische Topologie. Da diese das Hausdorffsche Umgebungsaxiom (D) erfüllt, genügt auch die schwächere C -Modifikation v diesem Axiom.

Damit wir auch den zweiten Teil dieses Satzes beweisen, führen wir in der Ebene eine waagerechte und eine lotrechte Topologie ein; in der waagerechten (lotrechten) Topologie wird als die Umgebung des Punktes $(a, b) \in R_2$ jede waagerechte (lotrechte) Strecke mit dem Mittelpunkt (a, b) bezeichnet. Die Länge der Strecke nennen wir die Länge der Umgebung. Wir konstruieren in der Ebene R_2 eine in der Topologie u abgeschlossene Menge F . Es sei $\{r^m\}$ die Folge aller rationalen Zahlen. Die Menge $F = \sum_{m=1}^{\infty} \left(r^m, \frac{1}{m}\right)$ ist in der

⁶⁾ Dieser Satz ist bekannt; man benutzt ihn in der Theorie der Baireschen Funktionen. Vgl. z. B. H. Hahn, Theorie der reellen Funktionen, S. 319. Berlin 1921.

⁷⁾ Zu jeder Topologie u gibt es eine kleinste Ordinalzahl α derart, daß $u^\alpha M = u^{\alpha+1} M$ für jede Untermenge M . Man nennt die Topologie u^α , die das Hausdorffsche Umgebungsaxiom (C) erfüllt, die C -Modifikation der Topologie u . Vgl. E. Čech, Topologické prostoty, Čas. pro příst. mat. a fys. 66 (1937), S. D 235.

⁸⁾ Dieser Satz gibt die Lösung eines Problems von E. Čech im topologischen Seminar vom 14. November 1939.

Kreuztopologie abgeschlossen. In der Tat, es gibt zu jedem Punkte $(a, b) \in R_2 - F$ eine waagerechte und eine lotrechte Umgebung, die keinen Punkt der Menge F enthält. Die Menge $R_2 - F$ ist also offen in der Topologie u und die Menge F abgeschlossen. Daher ist $u^0 F = u^1 F$, so daß $u^{n+1} F = u^n F$. Die Menge F ist also auch in der C -Modifikation v abgeschlossen. Es sei nun $G \supset F$ eine in der Topologie v offene Menge und es sei $V((a, 0))$ eine Umgebung des Punktes $(a, 0) \in R_2$ in der Topologie v . Da diese Topologie das Umgebungsaxiom (C) erfüllt, kann man voraussetzen, daß die Umgebung $V((a, 0))$ eine offene Menge ist. Die C -Modifikation der Kreuztopologie ist stärker als die Kreuztopologie und diese ist wieder stärker als die waagerechte und auch die lotrechte Topologie. Es ist daher eine Kreuzumgebung

$$U_p((a, 0)) = E\left[|x - a| < \frac{1}{p}, y = 0\right] + E\left[x = a, |y| < \frac{1}{p}\right] \subset V((a, 0))$$

vorhanden und zu jedem Punkte $(r^m, \frac{1}{m}) \in F$ gibt es eine waagerechte Umgebung $O((r^m, \frac{1}{m}))$, die ein Teil der Menge G ist.

Wir bezeichnen mit dem Symbol $K(x, y)$, wo $y \neq 0$ ist, ein Rechteck mit dem Mittelpunkt $(x, 0)$, dessen Seiten waagerecht und lotrecht sind und dessen eine Seite durch den Punkt (x, y) geht. Durch vollständige Induktion konstruieren wir eine Folge von Rechtecken folgendermaßen: Haben wir die Punkte $(x^i, y^i) \in F$ für $i = 1, 2, \dots, m-1$, und die Rechtecke $K(x^i, y^i)$, die im Innern des Rechtecks mit den Ecken $(a, -\frac{1}{p})$, $(a + \frac{1}{p}, -\frac{1}{p})$, $(a, \frac{1}{p})$, $(a + \frac{1}{p}, \frac{1}{p})$ liegen, schon gewählt, so bestimmen wir innerhalb des Rechtecks $K(x^{m-1}, y^{m-1})$ einen Punkt $(x^m, y^m) \in F$ und ein Rechteck $K(x^m, y^m)$ von folgenden Eigenschaften:

$$[m \ 1] \quad K(x^m, y^m) \subset K(x^{m-1}, y^{m-1}).$$

[m 2] Die Seite des Rechtecks $K(x^m, y^m)$, die durch den Punkt (x^m, y^m) geht, ist ein Teil der waagerechten Umgebung $O(x^m, y^m)$, die wir zu diesem Punkte bestimmt haben.

Nach dem Cantorschen Durchschnittssatz gibt es einen Punkt $(c, d) \in \bigcap_{n=1}^{\infty} K(x^n, y^n)$. Offenbar ist $a < c < a + \frac{1}{p}$. Es ist $\lim_{n \rightarrow \infty} y^n = 0$; denn die Zahlenfolge $\{y^n\}$ ist eine Teilfolge der Nullfolge $\{\frac{1}{n}\}$. Daher gilt: $d = 0$. In jeder lotrechten Umgebung des Punktes $(c, 0)$ liegen fast alle Punkte (c, y^n) . Da $(c, 0) \in U_p((a, 0)) \subset V((a, 0))$, so ist $(c, 0) \in V((a, 0))$. Weil die Umgebung $V((a, 0))$ eine offene Menge in der Topologie v und daher auch eine offene Menge in der schwächeren lotrechten Topologie ist, so liegen fast alle Punkte (c, y^n) in der Umgebung $V((a, 0))$, so daß $V((a, 0)) \cdot G \neq \emptyset$. Der Raum R_2 mit der Topologie v ist also nicht regulär.

Die C -Modifikation v der Kreuztopologie u ist effektiv stärker^{*)} als die Kreuztopologie u . In der Tat, diese zwei Topologien sind nicht gleichwertig, denn die C -Modifikation v erfüllt das Hausdorffsche Umgebungsaxiom (C), die Topologie u jedoch nicht.

2.

Es sei \mathfrak{S} irgendein nichtleeres System reeller Funktionen, die im Bereich $P \neq \emptyset$ definiert sind. Es sei $a \in P$, es sei ε eine beliebige positive Zahl und es sei $f \in \mathfrak{S}$. Unter der Umgebung $W(a; f, \varepsilon) \subset P$ des Punktes a verstehen wir die Menge aller derartigen Punkte $x \in P$, daß $|f(x) - f(a)| < \varepsilon$. Da immer $a \in W(a; f, \varepsilon)$, so ist das Hausdorffsche Umgebungsaxiom (A) erfüllt; daher ist der Bereich P ein topologischer Raum. Seine Topologie wird in dem Bereich P durch das Funktionensystem \mathfrak{S} induziert. Wir nennen sie daher die *Induktion des Systems* \mathfrak{S} oder meist einfacher die *Induktion*. Der Raum P heißt *Induktionsraum*; manchmal bezeichnen wir ihn mit (P, \mathfrak{S}) . Jede Funktion $f \in \mathfrak{S}$ ist im Raume (P, \mathfrak{S}) stetig, denn $W(a; f, \varepsilon)$ ist ja eine Umgebung von a .

Die Induktion genügt dem Hausdorffschen Umgebungsaxiom (C); ist nämlich $W(a; f, \varepsilon)$ eine Umgebung des Punktes a und $b \in W(a; f, \varepsilon)$, dann ist $W(b; f, \varepsilon - |f(b) - f(a)|) \subset W(a; f, \varepsilon)$. Es gilt der folgende Satz, den der Leser leicht beweist: *Der Induktionsraum (P, \mathfrak{S}) genügt dem Umgebungsaxiom (B) dann und nur dann, wenn die Summe zweier stetiger Funktionen eine stetige Funktion ist.*

Es sei eine Menge $P \neq \emptyset$ gegeben. Es sei \mathfrak{S} ein nichtleeres System reeller Funktionen, deren Bereich die Menge P ist. Wir können in der Menge P eine L -Topologie durch die folgende Definition der konvergenten Punktfolgen einführen: Die Punktfolge $\{z^m\}$, $z^m \in P$, konvergiert gegen den Punkt $z \in P$, falls $z^m = z$ für jedes $m = 1, 2, \dots$, oder $z^p \neq z^q$ für $p \neq q$; $p, q = 1, 2, \dots$, und der Punkt z der einzige derartige Punkt ist, daß $\lim_{m \rightarrow \infty} f(z^m) = f(z)$ für jede Funktion $f \in \mathfrak{S}$. Das System \mathfrak{S} induziert in der Menge P eine L -Topologie, die wir *L -Induktion des Systems* \mathfrak{S} oder einfacher *L -Induktion* nennen. Man beweist leicht den

Satz 2. *Die L -Induktion ist schwächer als die Induktion und die beiden Topologien sind dann und nur dann gleichwertig, falls die Induktion eine L -Topologie ist.*

Beweis. Es sei $W(a; f, \varepsilon)$ eine Umgebung eines Punktes $a \in P$. Konvergiert die Punktfolge $\{z^m\}$, $z^m \in P$, in der L -Induktion gegen den Punkt a ,

^{*)} Die Topologie u_2 ist effektiv stärker als die Topologie u_1 , wenn die Topologie u_2 stärker als die Topologie u_1 ist, aber beide Topologien nicht gleichwertig sind.

so ist $\lim_{m \rightarrow \infty} f(z^m) = f(a)$; es gibt also einen Index m_0 derart, daß $z^m \in W(a; f, \varepsilon)$ für alle $m > m_0$. Die Menge $W(a; f, \varepsilon)$ ist daher eine Umgebung des Punktes a in der L -Induktion.

Es sei nun die Induktion des Systems \mathfrak{S} eine L -Topologie. Wir beweisen, daß diese Induktion schwächer als die L -Induktion des Systems \mathfrak{S} ist. Es sei also $\{t^m\}$, $t^m \in P$, eine Punktfolge, die in der Induktion gegen den Punkt $b \in P$ konvergiert. Es ist zu beweisen¹⁰⁾, daß diese Punktfolge auch in der L -Induktion gegen den Punkt b konvergiert. Da jede Funktion $f \in \mathfrak{S}$ in dem Raume (P, \mathfrak{S}) stetig ist, so ist $\lim_{m \rightarrow \infty} f(t^m) = f(b)$. Ist auch $\lim_{m \rightarrow \infty} f(t^m) = f(b')$, $b' \in P$, für jede Funktion $f \in \mathfrak{S}$, so liegen in jeder Umgebung $W(b'; f, \varepsilon)$ des Punktes b' fast alle Punkte t^m , so daß die Punktfolge $\{t^m\}$ gegen den Punkt b' konvergiert. Es ist also $b = b'$ und b ist daher der einzige Punkt mit der Eigenschaft, daß $\lim_{m \rightarrow \infty} f(t^m) = f(b)$ für jede Funktion $f \in \mathfrak{S}$.

Es seien zwei natürliche Zahlen $k \leq n$ gegeben. Eine reelle Funktion $f(x_1, x_2, \dots, x_n)$ von n reellen Veränderlichen heißt nach H. Hahn¹¹⁾ *partiell stetig nach x_k im Punkte (a_1, a_2, \dots, a_n)* , wenn die Funktion $f(a_1, \dots, x_k, \dots, a_n)$ einer reellen Veränderlichen x_k im gewöhnlichen Sinne für $x_k = a_k$ stetig ist. Eine Funktion, die partiell stetig nach x_k in jedem Punkte des Bereiches ist, heißt *partiell stetig nach x_k* .

Es sei $\mathfrak{S}(x_k)$ das System aller partiell stetigen Funktionen nach x_k , $k = 1, 2$, in der Ebene R_2 . Die Induktion dieses Systems ist die metrische Topologie, und zwar: für $k = 1$ die waagerechte und für $k = 2$ die lotrechte Topologie.

Eine reelle Funktion $f(x_1, x_2, \dots, x_n)$ von n reellen Veränderlichen heißt *partiell stetig*¹¹⁾ [im Punkte (a_1, a_2, \dots, a_n)], wenn sie [in dem Punkte (a_1, a_2, \dots, a_n)] partiell stetig nach x_k , $k = 1, 2, \dots, n$, ist. Eine partiell stetige Funktion ist nicht immer im gewöhnlichen Sinne stetig. H. Hahn gibt folgendes Beispiel einer derartigen Funktion in der Euklidischen Ebene

$$f(x, y) = \frac{xy}{x^2 + y^2} \quad \text{für } (x, y) \neq (0, 0) \\ f(0, 0) = 0.$$

Der Leser überzeugt sich leicht davon, daß die L -Induktion *partiell stetiger Funktionen in der Ebene die Kreuztopologie u ist*.

Nach einem bekannten Satz¹²⁾ ist der Hausdorffsche Induktionsraum vollständig regulär und deshalb auch regulär. Die Induktion w ist also vermöge des Satzes 1 von der C -Modifikation v der L -Induktion u verschieden.

¹⁰⁾ Vgl. J. Novák, Sur les espaces (L) et sur les produits cartésiens (L) . Publ. de la fac. des sciences Brunn, Nr. 273 (1939), S. 11.

¹¹⁾ H. Hahn, Reelle Funktionen, S. 327. Leipzig 1932.

¹²⁾ Vgl. E. Čech, On bicompact spaces. Annals of Mathematics 38 (1937), S. 826.

Es sei \mathfrak{S} ein System aller partiell stetigen Funktionen von zwei reellen Veränderlichen. Die Summe von zwei partiell stetigen Funktionen ist wieder eine partiell stetige Funktion. Die Induktion des Systems \mathfrak{S} , die wir mit w bezeichnen wollen, erfüllt daher das Hausdorffsche Umgebungsaxiom (B). Da jede reelle Funktion, die im Bereich R_2 im gewöhnlichen Sinne stetig ist, auch partiell stetig ist, enthält das System \mathfrak{S} alle Funktionen, die im Euklidischen Raume R_2 stetig sind. Die Induktion w ist daher schwächer als die Euklidische Topologie. Die Topologie w erfüllt also alle vier Hausdorffschen Umgebungsaxiome.

Satz 3. *Der Induktionsraum (R_2, \mathfrak{S}) ist vollständig regulär, jedoch nicht normal¹³⁾.*

Wir beweisen zunächst den

Hilfssatz. *Jede Menge $H \subset (R_2, \mathfrak{S})$, die keinen Punkt $(x, y) \in R_2, x \neq y$, enthält, ist in dem Raume (R_2, \mathfrak{S}) abgeschlossen.*

Beweis. Es sei $(a, b) \in R_2 - H$. Es ist das Vorhandensein einer Umgebung $W((a, b); f, \varepsilon) \subset R_2$ zu beweisen, so daß $f \in \mathfrak{S}$, $\varepsilon > 0$ und $H \cdot W((a, b); f, \varepsilon) = \emptyset$. Ist $a \neq b$, so setzen wir $f(x, y) = (x - a)^2 + (y - b)^2$ und $\varepsilon = \frac{1}{4}(b - a)^2$. Ist jedoch $a = b$, dann definieren wir

$$f(x, y) = \frac{(x - a)(y - b)}{(x - a)^2 + (y - b)^2} \quad \text{für } (x, y) \neq (a, b),$$

$$f(a, b) = 0.$$

Da der erste Fall leicht ersichtlich ist, genügt es den zweiten zu beweisen. Die Funktion $f(x, y)$ ist partiell stetig, denn die Funktionen einer reellen Veränderlichen $f(x_0, y)$ und $f(x, y_0)$ sind stetig. Da $f(x, x) = \frac{1}{2}$ für jeden Punkt $(x, x) \in H$, so ist $H \cdot W((a, b); f, \varepsilon) = \emptyset$ für jede Zahl ε , die die Ungleichung $0 < \varepsilon \leq \frac{1}{2}$ erfüllt, w. z. b. w.

Es sei F_1 die Menge aller Punkte $(x, x) \in R_2$ mit irrationalem x und F_2 die Menge aller Punkte $(x, x) \in R_2$ mit rationalem x . Diese zwei Mengen sind disjunkt und nach dem Hilfssatz abgeschlossen. Es gibt aber keine Funktion $g(x, y)$ mit $g(F_1) = 0$, $g(F_2) = 1$, die im Induktionsraume (R_2, \mathfrak{S}) stetig wäre. Dies beweisen wir folgendermaßen: Es sei $g(x, y)$ eine im Raume (R_2, \mathfrak{S}) stetige Funktion mit $g(F_1) = 0$. Es sei ε eine beliebige positive Zahl. Die Menge $G = E[|g(x, y)| < \varepsilon, (x, y) \in R_2]$ ist in dem Raume (R_2, \mathfrak{S}) offen. Es ist $F_1 \subset G$. Da die lotrechte Topologie schwächer ist als die Induktion w , gibt es zu jedem Punkte $(a, a) \in F_1$ eine lotrechte Umgebung

¹³⁾ Dieser Raum ist sogar in keinem Punkte normal. Ich habe ein Beispiel eines solchen Raumes schon früher gegeben. Vgl. Zwei Bemerkungen zum Bernsteinschen Ultrakontinuum. Čas. pro pěst. mat. a fys. 68 (1939), S. 154.

$O_m((a, a)) \subset G$ von der Länge $\frac{1}{m}$ derart, daß $|g(x, y)| < \varepsilon$ für jeden Punkt $(x, y) \in O_m((a, a))$. Wir teilen alle irrationalen Zahlen nach folgender Regel in disjunkte Klassen K_j , $j = 1, 2, \dots$, ein: Die irrationale Zahl a reihen wir in die Klasse K_j dann und nur dann ein, wenn die zugeordnete Umgebung des Punktes $(a, a) \in F_1$ die Länge $\frac{1}{j}$ hat. Dieses Klassensystem ist offenbar disjunkt und abzählbar. Jede irrationale Zahl liegt in einer Klasse. Es gibt eine rationale Zahl c und einen Index p , so daß die Zahl c in der Euklidischen Hülle \bar{K}_p der Zahlenmenge K_p liegt. Im anderen Falle wäre $\sum_{j=1}^{\infty} \bar{K}_j$ die Menge aller irrationalen Zahlen, also ein F_σ , was bekanntlich unmöglich ist. Wir können annehmen, daß die Klasse K_p unabzählbar ist.

Da die dem Punkte $(x_0, x_0) \in K_p$ zugeordnete Umgebung die Menge $O_p((x_0, x_0)) = E[|y - x_0| < \frac{1}{p}, (x_0, y) \in R_2]$ ist, so enthält jede waagerechte Umgebung des Punktes $(c, c) \in F_2$ (und deshalb auch jede Umgebung dieses Punktes in der stärkeren Topologie w) solche Punkte $(x, y) \in R_2$, daß $|g(x, y)| < \varepsilon$. Da die Funktion $g(x, y)$ in der Induktion w im Punkte (c, c) stetig ist, so gilt: $|g(c, c)| \leq \varepsilon$. Weil die Zahl $\varepsilon > 0$ beliebig gewählt wurde, so ist $g(c, c) = 0$. Da $(c, c) \in F_2$, ist also nicht $g(x, y) = 1$ für jeden Punkt $(x, y) \in F_2$, w. z. b. w.

Um die Beziehung der Induktion w zu der Euklidischen Topologie näher zu bestimmen, schieben wir zwischen diese zwei Topologien eine Hilfsinduktion des Funktionensystems \mathfrak{P} ein, die wir mit w^* bezeichnen. In das System \mathfrak{P} sei jede Funktion folgender Form eingereiht:

$$\varphi(x_1, x_2) = (x_1 - a_1)^2 + (x_2 - a_2)^2$$

oder

$$\varphi(x_1, x_2) = \frac{|x_1 - a_1| \cdot |x_2 - a_2|}{(x_1 - a_1)^2 + (x_2 - a_2)^2} \quad \text{für } x_k \neq a_k, k = 1, 2,$$

$$\varphi(a_1, a_2) = 0$$

oder endlich

$$\varphi(x_1, x_2) + \varphi(x_1, x_2),$$

wobei a_k , $k = 1, 2$, alle reellen Zahlen durchläuft.

Satz 4. Die Induktion w^* ist effektiv schwächer als die Euklidische Topologie in R_2 und effektiv stärker als die Induktion w .

Beweis. Da das System aller Funktionen $\varphi(x_1, x_2) = (x_1 - a_1)^2 + (x_2 - a_2)^2$ in R_2 die Euklidische Topologie induziert und weil die Funktion $\varphi(x_1, x_2)$, die oben definiert ist, nicht im gewöhnlichen Sinne stetig ist, so enthält jede Euklidische Umgebung eines Punktes eine Umgebung desselben Punktes in der Topologie w^* , aber nicht umgekehrt. Die Induktion w^* ist daher effektiv schwächer als die Euklidische Topologie in der Ebene.

Um den zweiten Teil des Satzes zu beweisen, betrachten wir das Mengensystem

$$U_1([a_i]) \supset U_2([a_i]) \supset \dots \supset U_m([a_i]) \supset \dots,$$

wo $[a_i] = (a_1, a_2) \in R_2$ und $U_m([a_i])$ die Menge aller Punkte $(x_1, x_2) = [x_i] \in R_2$ ist, die gleichzeitig folgende Beziehungen befriedigen:

$$(x_1 - a_1)^2 + (x_2 - a_2)^2 < \frac{1}{m} \quad \text{und} \quad \frac{|x_1 - a_1| \cdot |x_2 - a_2|}{(x_1 - a_1)^2 + (x_2 - a_2)^2} < \frac{1}{m}.$$

Wir beweisen, daß die Induktion w^* und die Topologie, die jedem Punkte $[a_i] \in R_2$ als Umgebungen die Mengen $U_m([a_i])$ zuordnet, gleichwertig sind.

Es sei $W^*([a_i]; h, \varepsilon)$, wo $h \in \mathfrak{P}$ und $\varepsilon > 0$, eine Umgebung des Punktes $[a_i]$ in dem Raume (R_2, \mathfrak{P}) . Ist die Funktion $h(x_1, x_2)$ im Punkte $[a_i]$ im gewöhnlichen Sinne stetig, so gibt es einen solchen Index p , daß $U_p([a_i]) \subset W^*([a_i]; h, \varepsilon)$; befriedigt die Zahl p die Bedingung $\frac{1}{p} < \frac{\varepsilon}{2}$, so gilt diese Inklusion auch dann, wenn die Funktion $h(x_1, x_2)$ in dem Punkte $[a_i]$ nicht im gewöhnlichen Sinne stetig ist.

Es sei nun die Menge $U_m([a_i])$ und eine Zahl $\varepsilon > 0$ mit $\varepsilon < \frac{1}{m}$ gegeben. Dann gilt

$$W^*([a_i]; \varphi(x_1, x_2) + \psi(x_1, x_2), \varepsilon) \subset U_m([a_i]),$$

wobei die Funktionen φ und ψ von der oben angegebenen Form sind, w. z. b. w.

Aus der Gleichwertigkeit beider Topologien folgt, daß der Induktionsraum (R_2, \mathfrak{P}) ein Hausdorffscher Raum ist. Da das erste Hausdorffsche Abzählbarkeitsaxiom gilt, ist die Hilfsinduktion w^* eine L -Topologie. Dagegen ist die Induktion w keine L -Topologie, denn anderenfalls wäre die Induktion w nach dem Satz 2 eine L -Induktion, also gleichwertig mit der Kreuztopologie u , was jedoch nicht der Fall ist. Da die Funktionen φ , ψ und $\varphi + \psi$ partiell stetig sind, ist auch der zweite Teil des Satzes bewiesen.

Alle Betrachtungen über die Induktion partiell stetiger Funktionen von zwei reellen Veränderlichen lassen sich ohne Schwierigkeiten auf Funktionen einer beliebigen endlichen Anzahl von unabhängigen Variablen ausdehnen.

(Eingegangen am 20. 7. 1941.)

Störungstheorie der Spektralzerlegung. V.

Von

Franz Rellich in Dresden.

Ein selbstadjungierter Operator, der von einem Störungsparameter ε abhängt, habe im ungestörten Zustand, für $\varepsilon = 0$, ein diskretes Spektrum, d. h. sein Spektrum bestehe aus lauter Punkteigenwerten endlicher Vielfachheit $\lambda_1, \lambda_2, \dots$, die sich nicht im Endlichen häufen, das zugehörige vollständige System orthogonaler und normierter Eigenfunktionen sei $\varphi_1, \varphi_2, \dots$. Nach einem allgemeinen von mir bewiesenen Satz¹⁾ besitzt der gestörte Operator, vorausgesetzt, daß die Störung eine reguläre ist, zu jedem λ_n , φ_n ein $\lambda_n(\varepsilon)$, $\varphi_n(\varepsilon)$, das in der Umgebung von $\varepsilon = 0$ regulär analytisch von ε abhängt und für $\varepsilon = 0$ in λ_n, φ_n übergeht:

$$\begin{aligned}\lambda_n(\varepsilon) &= \lambda_n + \varepsilon \lambda_n^{(1)} + \varepsilon^2 \lambda_n^{(2)} + \dots \\ \varphi_n(\varepsilon) &= \varphi_n + \varepsilon \varphi_n^{(1)} + \varepsilon^2 \varphi_n^{(2)} + \dots\end{aligned}$$

Es entsteht die Frage: Bilden die so erklärten Eigenwerte $\lambda_n(\varepsilon)$ auch für $\varepsilon \neq 0$ das *gesamte* Spektrum des Operators, d. h. sind auch die *gestörten* Eigenfunktionen $\varphi_n(\varepsilon)$ vollständig? Die Frage ist in dieser Form noch nicht sinnvoll. Denn der Konvergenzradius von $\lambda_n(\varepsilon)$ bzw. von $\varphi_n(\varepsilon)$ wird im allgemeinen von n abhängen und kann mit wachsendem n gegen Null rücken; man kann also im allgemeinen bei festem $\varepsilon \neq 0$ auch für noch so kleines $|\varepsilon|$ gar nicht von der Gesamtheit der $\lambda_n(\varepsilon)$, $\varphi_n(\varepsilon)$ sprechen. Deshalb stellen wir die Frage so: Wenn sich alle $\lambda_n(\varepsilon)$ und $\varphi_n(\varepsilon)$ durch analytische Fortsetzung längs reeller ε in einem gemeinsamen Intervall $-\varrho < \varepsilon < \varrho$ erklären lassen, sind dann die so erklärten $\varphi_n(\varepsilon)$ für jedes ε dieses Intervalls vollständig? An zwei einfachen Beispielen zeigen wir, daß diese Frage zu *verneinen* ist.

1. Beispiel. Wir betrachten das Eigenwertproblem $u'' + \lambda u = 0$ mit den Randbedingungen $u(0) = 0$, $\varepsilon u'(1) + u(1) = 0$; es handelt sich physikalisch um die schwingende Saite, die links festgehalten und rechts durch eine Kraft $\frac{1}{\varepsilon} u(1)$ elastisch gekoppelt ist. Geht man zur entsprechenden Integralgleichung über, so ist deren Kern gegeben durch die Greensche Funktion $G(x, y; \varepsilon) = \frac{1}{2}(x + y) - \frac{1}{2}|x - y| - \frac{1}{1 + \varepsilon} xy$. Offenbar ist $G(x, y; \varepsilon)$ für jedes ε aus $|\varepsilon| < 1$ eine gleichmäßig für alle x, y aus $0 \leq x, y \leq 1$ konvergente

¹⁾ Satz 2 der ersten Mitteilung, Math. Annalen 113 (1936), S. 600–619.

Potenzreihe. Es handelt sich also um ein reguläres Störungsproblem²⁾. Die ungestörten Eigenwerte sind $\lambda_n = n^2 \pi^2$, die (normierten) Eigenfunktionen sind $\varphi_n = \sqrt{2} \sin(n\pi x)$. Um alle Eigenfunktionen des gestörten Problems zu finden, bemerke man zunächst, daß $\lambda = 0$ kein Eigenwert sein kann, wenn man sich auf ε beschränkt, für die $-1 < \varepsilon < 1$ ist; es sind dann alle Lösungen der Differentialgleichung mit $u(0) = 0$ von der Form $u(x) = a \sin(\sqrt{\lambda} x)$. Wegen $\varepsilon u'(1) + u(1) = 0$ muß λ der Gleichung $\operatorname{tg} \sqrt{\lambda} = -\varepsilon \sqrt{\lambda}$ genügen. Diese Gleichung besitzt eine Folge positiver Wurzeln $\lambda_n(\varepsilon) = v_n^2(\varepsilon)$ mit $\lambda_n(0) = n^2 \pi^2$; $n = 1, 2, \dots$. Die $v_n(\varepsilon)$ und damit auch die $\lambda_n(\varepsilon)$ sind in einer Umgebung von $\varepsilon = 0$ reguläre Potenzreihen von ε , die sich längs der ganzen reellen ε -Achse regulär analytisch fortsetzen lassen. Die gestörten (nicht normierten) Eigenfunktionen sind $\varphi_n(\varepsilon) = \sqrt{2} \sin[v_n(\varepsilon)x]$. Bei positivem ε sind diese vollständig, weil die Gleichung $\operatorname{tg} \sqrt{\lambda} = -\varepsilon \sqrt{\lambda}$ dann außer $\lambda_n = v_n^2(\varepsilon)$ keine reellen Wurzeln besitzt. Bei negativem ε aber sind sie nicht vollständig. Denn bei negativem ε besitzt die Gleichung $\operatorname{tg} \sqrt{\lambda} = -\varepsilon \sqrt{\lambda}$ oder $\operatorname{tg} \sqrt{-\lambda} = -\varepsilon \sqrt{-\lambda}$ (außer der uninteressanten Wurzel $\lambda = 0$) eine negative Wurzel $\lambda^*(\varepsilon) = -v^2(\varepsilon)$, für die $\lim_{\varepsilon \rightarrow 0} \lambda^*(\varepsilon) = -\infty$ gilt. Die zugehörige (nicht normierte) Eigenfunktion ist $\varphi^*(\varepsilon) = \sin[v^*(\varepsilon)x]$. Neben dem regulären Spektrum $\lambda_1(\varepsilon), \lambda_2(\varepsilon), \dots$ gibt es also noch den nichtregulären Eigenwert $\lambda^*(\varepsilon)$, der mit $\varepsilon \rightarrow 0, \varepsilon < 0$, gegen $-\infty$ rückt. Bei diesem (regulären!) Störungsproblem ist es also sogar der tiefste Eigenwert, der nicht durch Störungsrechnung erfaßt werden kann³⁾. — Die hier auftretende Instabilität ist indessen nicht das Wesentliche. Man braucht nur von dem eben beschriebenen Differentialoperator zu dessen Quadrat überzugehen, d. h. von der schwingenden Saite zum schwingenden Stab. Dann erhält man das

2. Beispiel. $u^{(4)} - \lambda u = 0, u(0) = u''(0) = 0, \varepsilon u'(1) + u(1) = 0, \varepsilon u'''(1) + u''(1) = 0$. Die Eigenwerte sind bei positivem ε gegeben durch $\lambda_n(\varepsilon) = [v_n(\varepsilon)]^4$. Bei negativem ε tritt der nichtreguläre Eigenwert $\lambda^*(\varepsilon) = [v^*(\varepsilon)]^4$ hinzu. Die Eigenfunktionen sind dieselben wie eben. Hier ist also $\lim_{\varepsilon \rightarrow 0} \lambda^*(\varepsilon) = +\infty$ und die Instabilität entfällt.

Es ist die Absicht dieser Mitteilung, für selbstadjungierte Operatoren, die regulär von einem Störungsparameter abhängen und im ungestörten

²⁾ Satz 4 der ersten Mitteilung und Satz 1 der dritten Mitteilung. Math. Annalen 116 (1939), S. 555–570.

³⁾ Physikalisch ist das Auftreten dieses $\lambda^*(\varepsilon)$ bei negativem ε unmittelbar einleuchtend, weil bei negativem ε auf das rechte Saitenende die abstoßende Kraft $\frac{1}{\varepsilon} u(1)$ wirkt.

Zustand ein diskretes Spektrum besitzen, hinreichende Bedingungen dafür anzugeben, daß im gestörten Zustand keine nichtregulären Bestandteile im Spektrum auftreten, daß vielmehr der gestörte Operator ein reguläres diskretes Spektrum besitzt.

Ich lege dabei den Begriff „reguläres diskretes Spektrum“ folgendermaßen fest:

Definition 1. Ein Operator $A(\varepsilon)$ sei in einem Teilraum $\mathfrak{A}(\varepsilon)$ eines unendlichdimensionalen Hilbertschen Raumes symmetrisch für alle ε des Intervalles $\varrho_1 < \varepsilon < \varrho_2$. Wir sagen, er besitzt in diesem Intervall ein reguläres diskretes Spektrum, wenn es Funktionen $\lambda_1(\varepsilon), \lambda_2(\varepsilon), \dots$ und Elemente $\varphi_1(\varepsilon), \varphi_2(\varepsilon), \dots$ aus $\mathfrak{A}(\varepsilon)$, beide regulär in der Umgebung jedes ε aus diesem Intervall derart gibt, daß für jedes ε aus $\varrho_1 < \varepsilon < \varrho_2$ gilt: 1) $A(\varepsilon)\varphi_n(\varepsilon) = \lambda_n(\varepsilon)\varphi_n(\varepsilon)$; $n = 1, 2, \dots$; 2) $\varphi_1(\varepsilon), \varphi_2(\varepsilon), \dots$ bildet ein vollständiges, orthogonales und normiertes System. 3) $\lim_{n \rightarrow \infty} |\lambda_n(\varepsilon)| = \infty$.

Wenn ein Operator in diesem Sinne ein reguläres diskretes Spektrum besitzt, dann kann er keine nichtregulären Eigenwerte (wie die Operatoren in den Beispielen 1 und 2) haben, weil der Satz gilt: Es seien die Voraussetzungen der Definition 1 erfüllt und es sei $\mu(\varepsilon)$ in einer Umgebung von $\varepsilon_0 (\varrho_1 < \varepsilon_0 < \varrho_2)$ eine reguläre Funktion und Eigenwert von $A(\varepsilon)$ [es gibt also zu jedem ε dieser Umgebung ein normiertes $\psi(\varepsilon)$, so daß $A(\varepsilon)\psi(\varepsilon) = \mu(\varepsilon)\psi(\varepsilon)$ gilt; $\psi(\varepsilon)$ braucht aber nicht regulär analytisch von ε abzuhängen]; dann muß $\mu(\varepsilon)$ mit einer der Funktionen $\lambda_1(\varepsilon), \lambda_2(\varepsilon), \dots$ identisch sein. In der Tat muß jedenfalls für jedes ε' eines Intervalls $|\varepsilon - \varepsilon_0| \leq \delta$, $\delta > 0$, das ganz in $\varrho_1 < \varepsilon < \varrho_2$ enthalten und in dem $\mu(\varepsilon)$ regulär analytisch ist, $\mu(\varepsilon')$ gleich einer der Zahlen $\lambda_1(\varepsilon'), \lambda_2(\varepsilon'), \dots$ sein; das folgt aus der Vollständigkeit der $\varphi_1(\varepsilon'), \varphi_2(\varepsilon'), \dots$. Würde nun $\mu(\varepsilon)$ im Intervall $|\varepsilon - \varepsilon_0| \leq \delta$ mit jeder der Funktionen $\lambda_1(\varepsilon), \lambda_2(\varepsilon), \dots$ nur an endlich vielen Stellen übereinstimmen, dann würde es nur abzählbar viele ε' geben, für die $\mu(\varepsilon')$ mit einer der Zahlen $\lambda_1(\varepsilon'), \lambda_2(\varepsilon'), \dots$ übereinstimmt, es bestünde das Kontinuum $|\varepsilon - \varepsilon_0| \leq \delta$ aus abzählbar vielen Punkten. Also muß $\mu(\varepsilon)$ mit einem $\lambda_i(\varepsilon)$ an unendlich vielen Stellen des Intervalls $|\varepsilon - \varepsilon_0| \leq \delta$ übereinstimmen, es müssen also wegen der Analytizität $\mu(\varepsilon)$ und $\lambda_i(\varepsilon)$ identisch sein. — Aus dieser Bemerkung folgt übrigens, daß die Funktionen $\lambda_1(\varepsilon), \lambda_2(\varepsilon), \dots$ der Definition 1 bis auf ihre Reihenfolge eindeutig bestimmt sind; wir bezeichnen sie daher als das reguläre diskrete Spektrum von $A(\varepsilon)$.

Ein Hauptergebnis (Satz 4) dieser Arbeit ist es nun, daß ein reguläres diskretes Spektrum für Operatoren, die im ungestörten Zustand ein diskretes Spektrum haben, jedenfalls dann gesichert ist, wenn der Operator $A(\varepsilon)$ in einem festen, von ε unabhängigen Definitionsbereich \mathfrak{A} selbstadjungiert ist

und für jedes u dieses Definitionsbereiches $A(\varepsilon)u$ ein reguläres Element ist⁴⁾. Dieser Satz ist von großer Allgemeinheit; für die klassischen Differentialoperatoren etwa bedeutet er, daß immer ein reguläres diskretes Spektrum vorliegt, wenn der Störungsparameter nicht in den „Randbedingungen“ vorkommt.

Für halbbeschränkte Operatoren wird darüber hinaus (Satz 5) reguläres diskretes Spektrum für Operatoren nachgewiesen, deren zugehörige Form (das ist für Differentialoperatoren das Integral des zugehörigen Variationsproblems) in einem von ε unabhängigen Definitionsbereich erklärt ist. Das trifft z. B. zu für das Eigenwertproblem $u'' + \lambda u = 0$; $u(0) = 0$, $u'(1) + \varepsilon u(1) = 0$. Das zugehörige Variationsproblem heißt $\int_0^1 u'^2 dx + \varepsilon u^2(1) = \text{Min.}$ $\int_0^1 u^2 dx = 1$, $u(0) = 0$ und es ist gestellt in dem von ε unabhängigen Bereich aller in $0 \leq x \leq 1$ totalstetigen Funktionen $u(x)$ mit $\int_0^1 u'^2 dx < \infty$ und $u(0) = 0$. Die Greensche Funktion dieses (regulären) Störungsproblems heißt $K(x, y; \varepsilon) = \frac{1}{2}(x + y) - \frac{1}{2}|x - y| - \frac{\varepsilon}{1 + \varepsilon}xy$.

Ein Operator mit diskretem Spektrum besitzt eine Reziproke, unter Umständen (wenn Null Eigenwert ist) muß man vor der Reziprokenbildung von A zu $A + a$ übergehen. Diese Reziproke ist vollstetig und besitzt als Eigenwerte die reziproken Eigenwerte von A bzw. von $A + a$; sie hat also ein Spektrum mit dem einzigen Häufungspunkt Null. Man kann daher die oben für Operatoren mit diskretem Spektrum gestellte Frage auch für vollstetige Operatoren stellen. Für solche Operatoren formuliere ich die

Definition 1a. *Ein Operator $A(\varepsilon)$ besitzt in dem Intervall $\varrho_1 < \varepsilon < \varrho_2$ ein reguläres Spektrum mit dem festen Häufungspunkt Null, wenn alle Forderungen der Definition 1 erfüllt sind mit Ausnahme der Forderung 3, die ersetzt wird durch: „3a) $\lim_{n \rightarrow \infty} |\lambda_n(\varepsilon)| = 0$; jedes $\lambda_n(\varepsilon)$ ist im Intervall $\varrho_1 < \varepsilon < \varrho_2$ entweder identisch Null oder von Null verschieden und diejenigen unter den $\varphi_n(\varepsilon)$, deren zugehörigen Eigenwerte verschwinden, sind von ε unabhängig.“*

Der Integraloperator mit dem Kern $G(x, y; \varepsilon) = \frac{1}{2}(x + y) - \frac{1}{2}|x - y| - \frac{1}{1 + \varepsilon}xy$ ist ein Beispiel eines regulären, beschränkten und symmetrischen,

⁴⁾ Man beachte, daß der Operator $A(\varepsilon)$ in seinem Definitionsbereich selbstadjungiert (= hypermaximal) sein muß im Sinne der Spektraltheorie. Im Beispiel ist der Operator selbstadjungiert im Bereich aller in $0 \leq x \leq 1$ einmal stetig differenzierbaren Funktionen $u(x)$, für die $u'(x)$ noch totalstetig ist mit $\int_0^1 u'^2 dx < \infty$ und $u(0) = 0$, $\varepsilon u'(1) + u(1) = 0$. Offenbar ist dieser Definitionsbereich von ε abhängig.

sogar vollstetigen Operatoren mit einem nichtregulären Eigenwert, nämlich dem Eigenwert⁵⁾ $-\frac{1}{\nu^2(\varepsilon)}$, wo $\nu(\varepsilon)$ die oben für $\varepsilon < 0$ erklärte analytische Funktion ist. Man zeigt leicht, daß $-\frac{1}{\nu^2(\varepsilon)}$ nicht in eine Umgebung von $\varepsilon = 0$ als reguläre analytische Funktion fortgesetzt werden kann. Dieser Integraloperator besitzt also kein reguläres Spektrum mit dem festen Häufungspunkt Null. (Begründung entsprechend der Bemerkung im Anschluß an Definition 1.)

Satz 1 und Satz 2 geben Kriterien dafür, daß ein regulärer (beschränkter und symmetrischer) Operator ein reguläres Spektrum mit dem festen Häufungspunkt Null besitzt. Danach wird klar (Satz 3), warum die Integralgleichung mit dem in $-1 < \varepsilon < 1$ regulären Kern $G(x, y; \varepsilon) = \frac{1}{2}(x + y) - \frac{1}{2}|x - y| - \frac{1}{1+\varepsilon}xy$ einen nichtregulären Eigenwert besitzt, während das für den Kern $K(x, y; \varepsilon) = \frac{1}{2}(x + y) - \frac{1}{2}|x - y| - \frac{\varepsilon}{1+\varepsilon}xy$ nicht der Fall ist.

Im § 4 beweise ich einen Satz, der für das Verständnis des Vorangehenden wichtig ist, der aber auch unabhängig von der Störungstheorie Interesse verdient: Wenn ein Operator mit diskretem Spektrum in einem Definitionsbereich \mathfrak{A} selbstadjungiert ist, dann hat jeder in \mathfrak{A} oder in einem Teilraum von \mathfrak{A} selbstadjungierte Operator ein diskretes Spektrum. Ein entsprechender Satz gilt für vollstetige Operatoren.

Daß reguläre Operatoren nichtreguläre Eigenwerte besitzen können, ist zuerst von K. Friedrichs⁶⁾ gezeigt worden, allerdings an einem Operator, der im ungestörten Zustand ein rein kontinuierliches Spektrum besitzt. Solche Operatoren bleiben hier außer Betracht; ich behandle sie in einer folgenden Mitteilung.

§ 1.

Vollstetige Operatoren.

Die Bezeichnungen sind dieselben wie in den vorangehenden Mitteilungen. Insbesondere wird der zugrunde gelegte Hilbertsche Raum mit \mathfrak{H} bezeichnet; $x, y, u, v, \varphi, \psi, \dots$ bedeuten Elemente aus \mathfrak{H} ; (x, y) ist das innere Produkt, $|x| = \sqrt{(x, x)}$. Unter „Operator“ wird stets „linearer Operator“ verstanden. Ein Operator A , erklärt in einem dichten Teilraum \mathfrak{A} von \mathfrak{H} (\mathfrak{A} heißt der Definitionsbereich von A), heißt symmetrisch, wenn $(Au, v) = (u, Av)$ ist

⁵⁾ Unter Eigenwert einer Integralgleichung mit dem Kern $G(x, y)$ verstehen wir eine Zahl λ mit der $\int G(x, y) \varphi(y) dy = \lambda \varphi(x)$ gilt, $\varphi(x) \neq 0$. In der älteren Literatur wird im Gegensatz dazu $\frac{1}{\lambda}$ als Eigenwert bezeichnet.

⁶⁾ K. Friedrichs, Über die Spektralzerlegung eines Integraloperators. Math. Annalen 116 (1938), S. 249–272. Anmerkung ³⁾.

für alle u, v aus \mathfrak{A} . Er heißt selbstadjungiert, wenn er symmetrisch ist und wenn aus dem Bestehen der Gleichung $(Au, f) = (u, g)$ für alle u aus \mathfrak{A} und ein festes Paar f, g aus \mathfrak{H} folgt, daß f in \mathfrak{A} liegt (woraus sich dann $Af = g$ ergibt).

Den beiden Hauptsätzen dieses Paragraphen schicken wir zwei Hilfssätze voraus.

Hilfssatz 1. *Es sei im Intervall $-\varrho < \varepsilon < \varrho$ ein l -dimensionaler Projektionsoperator $P(\varepsilon)$ erklärt und in der Umgebung eines jeden ε aus diesem Intervall regulär⁷⁾. Dann gibt es l orthogonale normierte Elemente $\omega^1(\varepsilon), \dots, \omega^l(\varepsilon)$, regulär analytisch in der Umgebung⁸⁾ eines jeden ε aus $-\varrho < \varepsilon < \varrho$, welche für jedes ε dieses Intervalles die zu $P(\varepsilon)$ gehörige Mannigfaltigkeit $\mathfrak{M}(\varepsilon)$ aufspannen.*

Beweis. Seien x^1, x^2, \dots von ε unabhängige Elemente, die ganz \mathfrak{H} aufspannen. Dann spannen $P(\varepsilon)x^1, P(\varepsilon)x^2, \dots$ die Mannigfaltigkeit $\mathfrak{M}(\varepsilon)$ auf. Von diesen müssen für $\varepsilon = 0$ genau l linear unabhängig sein; es seien das die Elemente $\psi^1(\varepsilon) = P(\varepsilon)x^1, \dots, \psi^l(\varepsilon) = P(\varepsilon)x^l$. Die Elemente $\psi^1(\varepsilon), \dots, \psi^l(\varepsilon)$ sind für jedes ε aus dem Intervall $-\varrho < \varepsilon < \varrho$ erklärt und können in der Umgebung keines ε aus diesem Intervall identisch verschwinden, weil sonst das betreffende ψ auch für $\varepsilon = 0$ verschwände und dann $\psi^1(0), \dots, \psi^l(0)$ nicht mehr die l -dimensionale Mannigfaltigkeit $\mathfrak{M}(0)$ aufspannen könnten. Nun Orthogonalisierungsprozeß: 1. $\omega^1(\varepsilon) = \frac{\psi^1(\varepsilon)}{\sqrt{(\psi^1(\varepsilon), \psi^1(\varepsilon))}}$. Es ist

$\omega^1(\varepsilon)$ erklärt in $-\varrho < \varepsilon < \varrho$. Sei ε_0 eine Stelle mit $-\varrho < \varepsilon_0 < \varrho$. Dann ist $P(\varepsilon)$ in einer Umgebung von $\varepsilon = \varepsilon_0$ in eine Potenzreihe nach $\varepsilon - \varepsilon_0$ entwickelbar, also $\psi^1(\varepsilon) = \psi^1(\varepsilon - \varepsilon_0)^k + \psi^1_{k+1}(\varepsilon - \varepsilon_0)^{k+1} + \dots, \psi^1_k \neq 0$. Also auch $\omega^1(\varepsilon)$ regulär in der Umgebung von $\varepsilon = \varepsilon_0$. 2. $\varphi^2(\varepsilon) = \psi^2(\varepsilon) - c_{21}\omega^1(\varepsilon)$ mit $c_{21} = (\psi^2(\varepsilon), \omega^1(\varepsilon))$. Dann $(\varphi^2(\varepsilon), \omega^1(\varepsilon)) = 0$ und offenbar $\varphi^2(\varepsilon)$ regulär in der Umgebung eines jeden ε aus $-\varrho < \varepsilon < \varrho$. Nach dem eben Bewiesenen ist $\omega^2(\varepsilon) = \frac{\varphi^2(\varepsilon)}{\sqrt{(\varphi^2(\varepsilon), \varphi^2(\varepsilon))}}$ gleichfalls regulär in der Umgebung eines jeden ε aus $-\varrho < \varepsilon < \varrho$. So fortfahrend erhält man den vollen Satz.

Hilfssatz 2. *Es sei $R(\varepsilon)$ in $|\varepsilon| < \varrho$ beschränkt, symmetrisch und regulär. In der Umgebung jedes ε aus dem Intervall $|\varepsilon| < \varrho$ sei $\lambda(\varepsilon)$ eine reguläre Funktion von ε und isolierter Eigenwert endlicher (positiver) Vielfachheit von $R(\varepsilon)$. An irgendeiner Stelle ε_0 ($|\varepsilon_0| < \varrho$) sei $\lambda(\varepsilon_0)$ ein h -facher Eigenwert von $R(\varepsilon_0)$, so daß¹⁾ es h , in einer Umgebung von ε_0 reguläre Eigenwerte $\lambda^{(1)}(\varepsilon), \dots, \lambda^{(h)}(\varepsilon)$ gibt, die für $\varepsilon = \varepsilon_0$ in $\lambda(\varepsilon_0)$ übergehen; von diesen h Funktionen seien in einer Umgebung von $\varepsilon = \varepsilon_0$ genau l mit $\lambda(\varepsilon)$ identisch, also $l \leq h$. Dann gibt es l (aber nicht*

⁷⁾ Definition 3 und 3' der ersten Mitteilung.

⁸⁾ In dieser Arbeit bedeutet ε eine reelle Veränderliche, wenn nicht ausdrücklich das Gegenteil gesagt wird.

mehr) orthogonale, normierte Elemente $\omega^1(\varepsilon), \dots, \omega^l(\varepsilon)$, die in der Umgebung eines jeden ε aus dem Intervall $|\varepsilon| < \varrho$ regulär sind und dort die Eigenwertgleichung $R(\varepsilon)\omega^{(i)}(\varepsilon) = \lambda(\varepsilon)\omega^{(i)}(\varepsilon)$ erfüllen ($i = 1, 2, \dots, l$).

Beweis. I. Es soll zuerst gezeigt werden, daß $\lambda(\varepsilon)$ in $-\varrho < \varepsilon < \varrho$ überall mindestens die Vielfachheit l besitzt und daß es kein Teilintervall geben kann, in dem die Vielfachheit von $\lambda(\varepsilon)$ durchweg größer wäre als l . Wäre die erste Hälfte dieser Behauptung falsch, dann gäbe es ein δ , ohne Beschränkung der Allgemeinheit nehmen wir $\varepsilon_0 < \delta$ an, derart, daß $\lambda(\varepsilon)$ in $\varepsilon_0 \leq \varepsilon < \delta$ mindestens l -facher Eigenwert ist, während das nicht mehr gilt in einer noch so kleinen vollen Umgebung von $\varepsilon = \delta$. Die Zahl $\lambda(\delta)$ habe die Vielfachheit κ , also $0 < \kappa < \infty$. Da $\lambda(\delta)$ nach Voraussetzung ein isolierter Eigenwert ist, gibt es¹⁾ eine Umgebung von $\varepsilon = \delta$ derart, daß das Spektrum von $R(\varepsilon)$ dort aus genau κ Eigenwerten $\lambda^{(1)}(\varepsilon), \dots, \lambda^{(\kappa)}(\varepsilon)$ besteht. Da $\lambda(\varepsilon)$ links von δ ein mindestens l -facher Eigenwert ist, folgt $\kappa \geq l$ und es müssen von den $\lambda^{(1)}(\varepsilon), \dots, \lambda^{(\kappa)}(\varepsilon)$ jedenfalls l untereinander identisch und gleich $\lambda(\varepsilon)$ sein. Das heißt aber, daß $\lambda(\varepsilon)$ doch in einer vollen Umgebung von $\varepsilon = \delta$ mindestens l -facher Eigenwert ist; also $\lambda(\varepsilon)$ in $-\varrho < \varepsilon < \varrho$ mindestens l -fach.

Gäbe es ein Intervall $\eta_1 < \varepsilon < \eta_2$ mit $\eta_1 > -\varrho$, $\eta_2 < \varrho$, in dem durchweg $\lambda(\varepsilon)$ eine größere Vielfachheit hätte als l , so könnte dieses Intervall nicht ε_0 enthalten. Ohne Beschränkung der Allgemeinheit nehmen wir $\varepsilon_0 < \eta_1$ an. Es müßte dann ein kleinstes δ geben mit $\varepsilon_0 \leq \delta \leq \eta_1$ derart, daß $\lambda(\varepsilon)$ für $\delta < \varepsilon < \eta_2$ durchweg eine größere Vielfachheit hätte als l . Die Anwendung von Satz 2 der ersten Mitteilung¹⁾ würde dann wie eben lehren, daß $\lambda(\varepsilon)$ auch noch in einer vollen Umgebung von $\varepsilon = \delta$ durchweg eine größere Vielfachheit hätte als l im Widerspruch zur Annahme über δ .

II. In der Umgebung eines jeden ε aus dem Intervall $|\varepsilon| < \varrho$ gibt es¹⁾ genau l orthogonale normierte und regulär-analytische Elemente, die dort Eigenelemente von $R(\varepsilon)$ sind mit dem Eigenwert $\lambda(\varepsilon)$; sie spannen zusammen den l -dimensionalen Raum $\mathfrak{M}_l(\varepsilon)$ auf, dessen Projektionsoperator $P(\varepsilon)$ heiße. Damit ist zu jedem ε aus $-\varrho < \varepsilon < \varrho$ ein in einer Umgebung dieses ε regulärer Projektionsoperator $P(\varepsilon)$ erklärt. Alle diese Operatoren gehen auseinander durch analytische Fortsetzung längs der reellen ε -Achse hervor. Denn hat man zwei Operatoren $P_1(\varepsilon)$ und $P_2(\varepsilon)$ in zwei übergreifenden Umgebungen erklärt, so muß der Durchschnitt dieser beiden Umgebungen unendlich viele sich im Innern dieses Durchschnitts häufende Punkte ε enthalten, in denen, $\lambda(\varepsilon)$ genau die Vielfachheit l besitzt. In diesen Punkten stimmen $P_1(\varepsilon)$ und $P_2(\varepsilon)$ überein; also sind sie wegen der Analytizität identisch. Anwendung von Hilfssatz 1 liefert die $\omega^1(\varepsilon), \dots, \omega^l(\varepsilon)$ mit den behaupteten Eigenschaften.

Ein Operator $R(\varepsilon)$, der für $|\varepsilon| < \varrho$ beschränkt und regulär ist, läßt sich vermöge der Entwicklung $R(\varepsilon) = R_0 + \varepsilon R_1 + \dots$ sofort auch für jedes

komplexe ε aus dem Kreis $|\varepsilon| < \varrho$ erklären. Bedeutet η irgendeine komplexe Zahl mit $|\eta| < \varrho$, dann ist $R(\varepsilon)$ in eine Potenzreihe nach $\varepsilon - \eta$ entwickelbar: $R(\varepsilon) = R^{(0)}(\eta) + R^{(1)}(\eta)(\varepsilon - \eta) + R^{(2)}(\eta)(\varepsilon - \eta)^2 + \dots$. Es gilt dann

Hilfssatz 3. Es sei $R(\varepsilon)$ beschränkt und regulär für $|\varepsilon| < \varrho$. Außerdem $|R(\varepsilon)x| \leq M|R_0x|$ bzw. $|(R(\varepsilon)x, x)| \leq M(R_0x, x)$ für alle x aus \mathfrak{H} und alle komplexen ε mit $|\varepsilon| < \varrho$ (es ist $R_0 = R(0)$). Dann ist $|R^{(n)}(\varepsilon)x| \leq \frac{M}{d^n}|R_0x|$ bzw. $|(R^{(n)}(\varepsilon)x, x)| \leq \frac{M}{d^n}(R_0x, x)$ für $|\varepsilon| \leq \varrho_1 < \varrho$ mit geeignetem positiven d .

Beweis. Es seien x, y zwei feste Elemente aus \mathfrak{H} und $f(\varepsilon) = (R(\varepsilon)x, y)$, also $\frac{1}{n!} \frac{d^n f(\varepsilon)}{d\varepsilon^n} = (R^{(n)}(\varepsilon)x, y)$. Es ist $f(\varepsilon)$ eine gewöhnliche Funktion von ε , regulär analytisch in $|\varepsilon| < \varrho$ und, wenn zunächst $|R(\varepsilon)x| \leq M|R_0x|$ angenommen wird, $|f(\varepsilon)| \leq |R_0x||y| \cdot M$ für $|\varepsilon| < \varrho$. Also $\left| \frac{f^{(n)}(\varepsilon)}{n!} \right| \leq \frac{|R_0x| \cdot |y| \cdot M}{d^n}$ oder $|(R^{(n)}(\varepsilon)x, y)| \leq \frac{|R_0x||y|M}{d^n}$ für $\varepsilon \leq \varrho_1$. Setzt man $y = R^{(n)}(\varepsilon)x$, so ergibt sich $|R^{(n)}(\varepsilon)x| \leq \frac{|R_0x|M}{d^n}$ wie behauptet. Ist aber $|(R(\varepsilon)x, x)| \leq M(R_0x, x)$, dann setze man $g(\varepsilon) = (R(\varepsilon)x, x)$. Es ist $g(\varepsilon)$ regulär analytisch in $|\varepsilon| < \varrho$ und es gilt dort $|g(\varepsilon)| \leq M(R_0x, x)$. Also wird $\left| \frac{g^{(n)}(\varepsilon)}{n!} \right| \leq \frac{M(R_0x, x)}{d^n}$, also $|(R^{(n)}(\varepsilon)x, x)| \leq \frac{M}{d^n}(R_0x, x)$ für $\varepsilon \leq \varrho_1$.

Satz 1. In $R(\varepsilon) = R_0 + \varepsilon R_1 + \dots$ seien R_0, R_1, \dots in einem (unendlich dimensional) Hilbertschen Raum \mathfrak{H} symmetrische Operatoren. Es sei R_0 vollstetig und $|R_n x| \leq k^n |R_0 x|$. Dann ist $R(\varepsilon)$ beschränkt, symmetrisch und regulär für $|\varepsilon| < \frac{1}{k}$ und besitzt im Intervall $-\frac{1}{2k} < \varepsilon < \frac{1}{2k}$ ein reguläres Spektrum^{*)} mit dem festen Häufungspunkt Null.

Beweis. I. Da R_0 vollstetig ist, ist es auch beschränkt, also $|R_0 x| \leq c|x|$ und wegen $|R_n x| \leq k^n c|x|$ sind auch die Operatoren R_1, R_2, \dots beschränkt. Nach Definition 3 der ersten Mitteilung ist $R(\varepsilon)$ ein beschränkter, symmetrischer, regulärer Operator für $|\varepsilon| < \frac{1}{k}$. Wie bei gewöhnlichen Potenzreihen ergibt sich die Möglichkeit, $R(\varepsilon)$ auch für komplexe ε mit $|\varepsilon| < \frac{1}{k}$ zu erklären; es wird $|R(\varepsilon)x| \leq |R_0x|(1 + |\varepsilon|k + |\varepsilon|^2 k^2 + \dots) = \frac{1}{1 - |\varepsilon|k} |R_0x|$.

Für den ganzen folgenden Beweis seien q und q_1 feste Zahlen mit

$$(1) \quad 0 < q_1 < q < \frac{1}{2k}.$$

*) Vgl. Definition 1a der Einleitung.

Also

$$(2) \quad |R(\varepsilon)x| \leq 2|R_0x| \quad \text{für komplexes } \varepsilon \text{ mit } |\varepsilon| \leq q.$$

Umgekehrt:

$$\begin{aligned} |R(\varepsilon)x| &\geq |R_0x| - |R_0x| \cdot (|\varepsilon|k + |\varepsilon|^2 k^2 + \dots) \geq \left(1 - \frac{|\varepsilon|k}{1-|\varepsilon|k}\right) |R_0x| \\ &= \frac{1-2|\varepsilon|k}{1-|\varepsilon|k} |R_0x| \geq (1-2qk) |R_0x| \end{aligned}$$

oder

$$(3) \quad |R_0x| \leq \frac{1}{1-2qk} |R(\varepsilon)x|.$$

Nach Hilfssatz 3 folgt aus (2) wie bei gewöhnlichen Potenzreihen die Ungleichung $|R'(\varepsilon)x| \leq \frac{2}{d} |R_0x|$ und nach (3) daraus

$$(4) \quad |R'(\varepsilon)x| \leq \frac{2}{d(1-2qk)} |R(\varepsilon)x| \quad \text{für } |\varepsilon| \leq q_1.$$

(Wir schreiben $R'(\varepsilon)$ für $R^{(1)}(\varepsilon)$). Im folgenden sei ε wieder reell, $|\varepsilon| \leq q_1$.

II. Aus der Vollstetigkeit von R_0 folgt wegen (2) die Vollstetigkeit von $R(\varepsilon)$. Denn bedeutet $\{x\}$ irgendeine Menge von Elementen mit $|x| \leq 1$, dann läßt sich eine Folge x^1, x^2, \dots auswählen, für die $R_0 x^n$ konvergiert. Aus $|R_0(x^n - x^m)| \rightarrow 0$ für $n, m \rightarrow \infty$ folgt aber $|R(\varepsilon)(x^n - x^m)| \rightarrow \infty$ für $n, m \rightarrow \infty$; d. h. $R(\varepsilon)$ ist vollstetig. Es besitzt also $R(\varepsilon)$ ein reines Punktspektrum. Jeder Eigenwert von $R(\varepsilon)$, der nicht Null ist, hat endliche Vielfachheit und ist ein isolierter Eigenwert. Die Gesamtheit der Elemente x mit $R(\varepsilon)x = 0$, also der Eigenraum, der zum Eigenwert Null von $R(\varepsilon)$ gehört, ist von ε unabhängig. In der Tat folgt aus $R(\varepsilon)x = 0$ nach (3) auch $R_0x = 0$ und umgekehrt aus $R_0x = 0$ nach (2) auch $R(\varepsilon)x = 0$.

III. Es sei $\lambda(\varepsilon)$ in der Umgebung eines jeden ε aus dem Intervall $\varrho_1 < \varepsilon < \varrho_2$ ($-q_1 < \varrho_1 < \varrho_2 < +q_1$) regulär und Eigenwert von $R(\varepsilon)$. Es sei η ein Punkt des Intervalls $\varrho_1 < \varepsilon < \varrho_2$ mit $\lambda(\eta) \neq 0$ und $\varphi(\varepsilon)$ ein zugehöriges, in einer Umgebung von $\varepsilon = \eta$ reguläres normiertes Eigenelement¹⁰⁾ von $R(\varepsilon)$. Aus $R(\eta)\varphi(\eta) = \lambda(\eta)\varphi(\eta)$ folgt nach (2) sofort $|\lambda(\eta)| \leq 2|R_0\varphi(\eta)|$, also wegen $|R_0x| \leq c|x|$

$$(5) \quad |\lambda(\eta)| \leq 2c.$$

In einer Umgebung von $\varepsilon = \eta$ ist $\lambda(\varepsilon) = \lambda(\eta) + \lambda'(\eta)(\varepsilon - \eta) + \dots$; $\varphi(\varepsilon) = \varphi(\eta) + \varphi'(\eta)(\varepsilon - \eta) + \dots$; $R(\varepsilon) = R(\eta) + R'(\eta)(\varepsilon - \eta) + \dots$. Daraus folgt (formale Störungsrechnung!) $\lambda'(\eta) = (R'(\eta)\varphi(\eta), \varphi(\eta))$. Also

¹⁰⁾ Daß ein solches existiert, folgt aus Satz 2 der ersten Mitteilung.

$|\lambda'(\eta)| \leq |R'(\eta) \varphi(\eta)|$ und nach (4) weiter $|\lambda'(\eta)| \leq \frac{2}{d(1-2qk)} |R(\eta) \varphi(\eta)|$.
Man erhält so die Ungleichung¹¹⁾

$$(6) \quad |\lambda'(\eta)| \leq \frac{2}{d(1-2qk)} |\lambda(\eta)|.$$

IV. Es sei $-q_1 < \varepsilon_0 < q_1$ und $\lambda \neq 0$ ein Eigenwert von $R(\varepsilon_0)$. Es folgt¹⁾ in einer Umgebung von $\varepsilon = \varepsilon_0$ die Existenz eines regulären Eigenwertes $\lambda(\varepsilon)$ von $R(\varepsilon)$ mit $\lambda(\varepsilon_0) = \lambda$. Es soll gezeigt werden, daß dieses $\lambda(\varepsilon)$ sich regulär analytisch längs des ganzen Intervalles $-q_1 \leq \varepsilon \leq q_1$ fortsetzen läßt und dort nicht verschwindet. Wäre das nicht der Fall, dann gäbe es ein ϑ (wir nehmen ohne Beschränkung der Allgemeinheit an $\vartheta > \varepsilon_0$) mit $\vartheta \leq q_1$ derart, daß $\lambda(\varepsilon)$ in $\varepsilon_0 \leq \varepsilon < \vartheta$ regulär und von Null verschieden ausfällt, während $\lambda(\varepsilon)$ in $\varepsilon = \vartheta$ entweder nicht regulär ist oder verschwindet. Da $|\lambda(\varepsilon)|$ nach Ungleichung (5) im Intervall $\varepsilon_0 \leq \varepsilon < \vartheta$ beschränkt ist, sind nur zwei Fälle denkbar: entweder es gibt eine Folge ε_n mit $\varepsilon_0 \leq \varepsilon_n < \vartheta$, $\lim_{n \rightarrow \infty} \varepsilon_n = \vartheta$ und $\lim_{n \rightarrow \infty} \lambda(\varepsilon_n) = \mu \neq 0$; oder $\lambda(\varepsilon) \rightarrow 0$ für $\varepsilon \rightarrow \vartheta$, $\varepsilon_0 \leq \varepsilon < \vartheta$.

Im ersten Fall ist μ ein isolierter Eigenwert endlicher Vielfachheit von $R(\vartheta)$, vielleicht von der Vielfachheit Null, d. h. überhaupt kein Eigenwert. Da nämlich $R(\varepsilon)$ in der Umgebung von $\varepsilon = \vartheta$ regulär ist, gibt es h , in der Umgebung von $\varepsilon = \vartheta$ reguläre Eigenwerte $\mu_1(\varepsilon), \dots, \mu_h(\varepsilon)$ (es bedeutet h die Vielfachheit von μ), die für $\varepsilon = \vartheta$ in μ übergehen, und diese Eigenwerte stellen bei genügend kleinem $|\varepsilon - \vartheta|$ genau das Spektrum von $R(\varepsilon)$ in der Umgebung von $\lambda = \mu$ dar (wenn $h = 0$, dann ist das Spektrum von $R(\varepsilon)$ leer in einer Umgebung von $\lambda = \mu$). Da $\lim_{n \rightarrow \infty} \lambda(\varepsilon_n) = \mu$ ist, kann h nicht Null sein und es muß also wegen der Analytizität $\lambda_i(\varepsilon) \equiv \mu_i(\varepsilon)$ gelten, wobei i eine Zahl zwischen 1 und h ist; also $\lambda(\varepsilon)$ regulär für $\varepsilon = \vartheta$. Wir zeigen zweitens, daß nicht $\lim_{\varepsilon \rightarrow \vartheta} \lambda(\varepsilon) = 0$ sein kann. Es gilt nämlich in $\varepsilon_0 \leq \varepsilon < \vartheta$ nach (6) die

Ungleichung $|\lambda'(\varepsilon)| \leq \frac{2}{d(1-2qk)} |\lambda(\varepsilon)|$. Es ist also $\lambda(\varepsilon)$ in $\varepsilon_0 \leq \varepsilon < \vartheta$ eine stetig differenzierbare, in $\varepsilon_0 \leq \varepsilon \leq \vartheta$ stetige Funktion; $\lambda(\vartheta) = 0$ und $|\lambda'(\varepsilon)| \leq C \cdot |\lambda(\varepsilon)|$ in $\varepsilon_0 \leq \varepsilon < \vartheta$. Nach der bekannten Schlußweise, mit der man den Eindeutigkeitssatz für gewöhnliche Differentialgleichungen erster Ordnung zeigt, ist $\lambda(\varepsilon) = 0$ in $\varepsilon_0 \leq \varepsilon \leq \vartheta$ im Gegensatz zu $\lambda(\varepsilon_0) = \lambda \neq 0$; damit ist auch im zweiten Fall ein Widerspruch hergeleitet.

Also läßt sich $\lambda(\varepsilon)$ über das ganze Intervall $-q_1 \leq \varepsilon \leq q_1$ oder, wegen der Willkür von q, q_1 , über das offene Intervall $-\frac{1}{2k} < \varepsilon < \frac{1}{2k}$ regulär analytisch fortsetzen und es ist dort $\lambda(\varepsilon) \neq 0$.

¹¹⁾ Die Heranziehung der Ableitung $\lambda'(\eta)$ mit Hilfe der Formel $\lambda'(\eta) = (R'(\eta) \varphi(\eta), \varphi(\eta))$ und ihre Verwendung in dem folgenden Schritt IV verdanke ich für einen wichtigen Spezialfall Herrn v. d. Waerden.

V. Es sei $\lambda_1, \lambda_2, \dots$ die Gesamtheit der Eigenwerte von $R_0 = R(0)$, mehrfache mehrfach hingeschrieben. Sei λ ein nichtverschwindender unter diesen Eigenwerten; er sei h -fach. Dann gibt es in einer Umgebung von $\varepsilon = 0$ genau h reguläre Eigenwerte $\lambda^{(1)}(\varepsilon), \dots, \lambda^{(h)}(\varepsilon)$ mit $\lambda^{(1)}(0) = \dots = \lambda^{(h)}(0) = \lambda$. Die Numerierung sei so vorgenommen, daß in einer Umgebung von $\varepsilon = 0$ gelte $\lambda^{(1)}(\varepsilon) \equiv \lambda^{(2)}(\varepsilon) \equiv \dots \equiv \lambda^{(l)}(\varepsilon); \lambda^{(l+1)}(\varepsilon) \equiv \lambda^{(l+2)}(\varepsilon) \equiv \dots \equiv \lambda^{(l_2)}(\varepsilon); \dots; \lambda^{(l_{r-1}+1)}(\varepsilon) \equiv \lambda^{(l_{r-1}+2)}(\varepsilon) \equiv \dots \equiv \lambda^{(l_r)}(\varepsilon); l_r = h$. Jede der Funktionen $\lambda^{(i_1)}(\varepsilon), \lambda^{(i_2)}(\varepsilon), \dots, \lambda^{(i_r)}(\varepsilon)$ läßt sich, wie in Schritt IV bewiesen, über das ganze Intervall $-\frac{1}{2k} < \varepsilon < \frac{1}{2k}$ regulär analytisch fortsetzen ohne zu verschwinden. Nach Hilfssatz 2 gibt es längs dieses Intervalles reguläre, orthogonale und normierte Eigenelemente $\omega^i(\varepsilon): R(\varepsilon)\omega^i(\varepsilon) = \lambda^{(i)}(\varepsilon)\omega^i(\varepsilon)$, $i = 1, 2, \dots, l_1, \dots, l_2, \dots, l_r = h$.

Indem man diese Konstruktion für jeden der verschiedenen Eigenwerte $\lambda_n \neq 0$ vornimmt und zu den so erhaltenen $\omega(\varepsilon)$ noch ein (endliches oder unendliches) System normierter Orthogonalfunktionen $\varphi^1, \varphi^2, \dots$ hinzufügt, die den Raum aller x mit $R_0 x = 0$ aufspannen, erhält man ein System von Elementen, das wir irgendwie abzählen und mit $\varphi^1(\varepsilon), \varphi^2(\varepsilon), \dots$ bezeichnen. Es ist $R(\varepsilon)\varphi^n(\varepsilon) = \lambda_n(\varepsilon)\varphi^n(\varepsilon)$, wobei $\lambda_n(0) = \lambda_n$ ist und die $\varphi^n(\varepsilon)$ mit $\lambda_n = 0$ von ε unabhängig sind; $\lambda_n(\varepsilon)$ und $\varphi^n(\varepsilon)$ regulär in jedem Punkt aus dem Intervall $-\frac{1}{2k} < \varepsilon < \frac{1}{2k}$.

VI. Bleibt zu zeigen, daß die $\varphi^1(\varepsilon), \varphi^2(\varepsilon), \dots$ für jedes ε aus $-\frac{1}{2k} < \varepsilon < \frac{1}{2k}$ vollständig sind. Für $\varepsilon = 0$ ist das nach Konstruktion der Fall. Wäre $\varphi^1(\varepsilon_0), \varphi^2(\varepsilon_0), \dots$ nicht vollständig mit $-\frac{1}{2k} < \varepsilon_0 < \frac{1}{2k}$, dann würde es ein Eigen-element $\psi \neq 0$ geben, das auf allen $\varphi^1(\varepsilon_0), \varphi^2(\varepsilon_0), \dots$ orthogonal wäre; $R(\varepsilon_0)\psi = \mu\psi$. Wäre $\mu = 0$, so wäre (Schritt II) auch $R_0\psi = 0$ und außerdem ψ orthogonal auf allen $\varphi^1, \varphi^2, \dots$, die ja unter den $\varphi^1(\varepsilon_0), \varphi^2(\varepsilon_0), \dots$ vorkommen. Das ist unmöglich, weil wegen $R_0\psi = 0$ gewiß $\psi = \sum_r (\psi, \varphi^r) \varphi^r$ ist, also $\psi = 0$ folgen würde. — Wenn aber $\mu \neq 0$, dann gibt es¹⁾ zunächst in einer Umgebung von $\varepsilon = \varepsilon_0$ mindestens einen regulären Eigenwert von $R(\varepsilon)$, der für $\varepsilon = \varepsilon_0$ in μ übergeht. Nach Schritt IV läßt sich dieser Eigenwert regulär analytisch über das ganze Intervall $-\frac{1}{2k} < \varepsilon < \frac{1}{2k}$ fortsetzen ohne zu verschwinden. Nach Hilfssatz 2 gibt es (mindestens) ein normiertes $\varphi(\varepsilon)$, regulär in jedem ε aus $-\frac{1}{2k} < \varepsilon < \frac{1}{2k}$, das Eigenelement zum Eigenwert $\mu(\varepsilon)$ ist: $R(\varepsilon)\varphi(\varepsilon) = \mu(\varepsilon)\varphi(\varepsilon)$. Es muß $\mu(0)$ mit einer nichtverschwindenden unter den Zahlen $\lambda_1(0), \lambda_2(0), \dots$ übereinstimmen, also¹⁾ $\mu(\varepsilon)$ mit einem der $\lambda_1(\varepsilon), \lambda_2(\varepsilon), \dots$ identisch sein, wir nehmen an $\mu(\varepsilon) \equiv \lambda_1(\varepsilon)$. Es sei $\lambda_1(\varepsilon) \equiv \lambda_2(\varepsilon) \equiv \dots \equiv \lambda_l(\varepsilon)$, aber $\lambda_r(\varepsilon) \neq \lambda_1(\varepsilon)$, wenn $r > l$. Dann wird in

einer Umgebung von $\varepsilon = 0$ offenbar $\varphi(\varepsilon) \equiv \sum_{i=1}^l (\varphi(\varepsilon), \varphi^i(\varepsilon)) \varphi^i(\varepsilon)$. Da aber $\varphi(\varepsilon)$ und die $\varphi^i(\varepsilon)$ sich über $-\frac{1}{2k} < \varepsilon < \frac{1}{2k}$ analytisch fortsetzen lassen, gilt $\varphi(\varepsilon_0) = \sum_{i=1}^l (\varphi(\varepsilon_0), \varphi^i(\varepsilon_0)) \varphi^i(\varepsilon_0)$. Da $\varphi(\varepsilon_0)$ auf allen $\varphi^1(\varepsilon_0), \varphi^2(\varepsilon_0), \dots$ orthogonal steht, folgt $\varphi(\varepsilon_0) = 0$ im Widerspruch zu $|\varphi(\varepsilon_0)| = 1$. Damit ist Satz 1 vollständig bewiesen.

Wenn man von dem ungestörten Operator $R_0 = R(0)$ weiß, daß $(R_0 x, x) \geq 0$ ist für alle x , dann kann man in Satz 1 die Voraussetzungen abschwächen und gelangt zu

Satz 2. In $R(\varepsilon) = R_0 + \varepsilon R_1 + \dots$ seien R_0, R_1, \dots in einem (unendlich dimensional)en Hilbertschen Raum \mathfrak{H} symmetrische Operatoren. Es sei R_0 vollstetig und $|(R_n x, x)| \leq k^n (R_0 x, x)$. Dann ist $R(\varepsilon)$ beschränkt, symmetrisch und regulär für $|\varepsilon| < \frac{1}{k}$ und besitzt im Intervall $-\frac{1}{2k} < \varepsilon < \frac{1}{2k}$ ein reguläres Spektrum mit dem festen Häufungspunkt Null.

Beweis. Für den ganzen Beweis seien q und q_1 feste Zahlen und zwar $0 < q_1 < q < \frac{1}{2k}$. Aus $|R_0 x| \leq c|x|$ folgt $|(R_n x, x)| \leq k^n \cdot c \cdot (x, x)$, $|R_n x| \leq k^n c|x|$. Also $R(\varepsilon) = R_0 + \varepsilon R_1 + \dots$ regulär in $|\varepsilon| < \frac{1}{k}$. Ferner $|(R(\varepsilon)x, x)| \leq (R_0 x, x)(1 + |\varepsilon|k + \dots)$ also

$$(7) \quad |(R(\varepsilon)x, x)| \leq 2(R_0 x, x) \text{ für komplexes } \varepsilon \text{ aus } |\varepsilon| \leq q.$$

Umgekehrt⁸⁾ $(R(\varepsilon)x, x) \geq (R_0 x, x)(1 - |\varepsilon| \cdot k - \dots) \geq (1 - 2qk)(R_0 x, x)$ oder

$$(8) \quad (R_0 x, x) \leq \frac{1}{1-2qk} (R(\varepsilon)x, x) \text{ für (reelles) } \varepsilon \text{ aus } |\varepsilon| \leq q.$$

Aus (7) folgt nach Hilfssatz 3

$$(8a) \quad |(R'(\varepsilon)x, x)| \leq \frac{2}{d} (R_0 x, x) \leq \frac{2}{d} \frac{1}{1-2qk} (R(\varepsilon)x, x) \quad |\varepsilon| \leq q_1.$$

Bedeutet $\lambda(\varepsilon)$ einen Eigenwert von $R(\varepsilon)$, der in jedem Punkt eines Intervalles, das noch ganz zu $|\varepsilon| \leq q_1$ gehört, regulär ist, dann ist $\lambda'(\varepsilon) = (R'(\varepsilon)\varphi(\varepsilon), \varphi(\varepsilon))$ und $R(\varepsilon)\varphi(\varepsilon) = \lambda(\varepsilon)\varphi(\varepsilon)$, $|\varphi(\varepsilon)| = 1$. Wegen (8a) wird $|\lambda'(\varepsilon)| \leq \frac{2}{d} \frac{1}{1-2qk} |\lambda(\varepsilon)|$ für $\varepsilon \leq q_1$ und wegen (7) wird $|\lambda(\varepsilon)| \leq 2c$. Schließlich folgt, daß die Gesamtheit aller x mit $R(\varepsilon)x = 0$ von ε unabhängig ist ($|\varepsilon| \leq q$). Aus $R_0 x = 0$ folgt nämlich nach (7) $(R(\varepsilon)x, x) = 0$ und da⁹⁾ nach (8) auch $(R(\varepsilon)u, u) \geq 0$ ist für alle u , daß $R(\varepsilon)x = 0$ ist; umgekehrt folgt aus $R(\varepsilon)x = 0$ nach (8) auch $(R_0 x, x) = 0$, d. h. $R_0 x = 0$. — Da R_0 vollstetig ist, läßt sich aus jeder Menge $\{x\}$ mit $(x, x) \leq 1$ eine Folge x^1, x^2, \dots auswählen mit konvergentem $R_0 x^n$. Aus (7) folgt für $-q \leq \varepsilon \leq q$ auch¹²⁾ $(R(\varepsilon)x, R(\varepsilon)x) \leq 4c(R_0 x, x) \leq 4c|R_0 x| \cdot |x|$.

Also $|R(\varepsilon)(x^n - x^m)|^2 \leq 4c \cdot |R_0(x^n - x^m)| \cdot 2 \rightarrow 0$ für $n, m \rightarrow \infty$; d. h. $R(\varepsilon)$ ist vollstetig. Damit sind die Überlegungen zu Ende geführt, die an Stelle der Schritte I—III im Beweis von Satz 1 treten. Die restlichen Schritte IV—VI können unverändert aus diesem Beweis genommen werden und Satz 2 ist bewiesen.

In Satz 1 bzw. in Satz 2 sind über die symmetrischen beschränkten Operatoren R_0, R_1, \dots die unendlich vielen Voraussetzungen

$$(9) \quad |R_n x| \leq k^n |R_0 x| \quad \text{bzw.} \quad |(R_n x, x)| \leq k^n (R_0 x, x); \quad n = 1, 2, \dots$$

gemacht worden, aus denen sich nebenbei die Regularität von $R(\varepsilon) = R_0 + \varepsilon R_1 + \dots$ in $|\varepsilon| < \frac{1}{k}$ ergab. Man kann die unendlich vielen Voraussetzungen (9) aber in folgender Weise (durch äquivalente Voraussetzungen) ersetzen. Man verlange: 1. $R(\varepsilon) = R_0 + \varepsilon R_1 + \dots$ in $|\varepsilon| < \varrho$ regulär beschränkt und⁹⁾ symmetrisch. 2. Für alle komplexen ε aus dem Kreis $|\varepsilon| < \varrho$ sei

$$(10) \quad |R(\varepsilon)x| \leq M |R_0 x| \quad \text{bzw.} \quad |(R(\varepsilon)x, x)| \leq M (R_0 x, x).$$

In der Tat folgt aus (10) wegen Hilfssatz 3 (es ist $R^{(n)}(0) = R_n$) sofort auch (9). Daß umgekehrt aus (9) auch (10) folgt, ist klar.

§ 2.

Integraloperatoren.

Die Sätze 1 und 2 lassen sich direkt auf Integraloperatoren anwenden und liefern dann unter Berücksichtigung der Bemerkung am Schluß des vorangehenden Paragraphen [Formel (10)] den

Satz 3. Es sei 1. $K(x, y; \varepsilon)$ in einer festen Umgebung von $\varepsilon = 0$ eine gleichmäßig für alle x, y mit $0 \leq x, y \leq 1$ konvergente Potenzreihe von ε mit Koeffizienten, die im Intervall $0 \leq x, y \leq 1$ stetige Funktionen von x, y sind.

2. $K(x, y; \varepsilon) = \overline{K(y, x; \varepsilon)}$ für reelle ε .

3. Es gebe eine Zahl M und eine feste komplexe Umgebung $|\varepsilon| < \varrho'$ von $\varepsilon = 0$ derart, daß für alle in $0 \leq x \leq 1$ stetigen (komplexwertigen) Funktionen $u(x)$ gilt

$$\text{entweder: } \int_0^1 \left| \int_0^1 K(x, y; \varepsilon) u(y) dy \right|^2 dx \leq M^2 \int_0^1 \left| \int_0^1 K(x, y; 0) u(y) dy \right|^2 dx$$

$$\text{oder: } \left| \int_0^1 \int_0^1 K(x, y; \varepsilon) u(x) \overline{u(y)} dx dy \right| \leq M \int_0^1 \int_0^1 K(x, y; 0) u(x) \overline{u(y)} dx dy.$$

¹²⁾ Wenn B und B' beschränkte, symmetrische Operatoren sind und wenn $|(Bu, u)| \leq k(B'u, u)$ und $|B'u| \leq b'|u|$ ist, dann folgt zunächst in bekannter Weise $|(Bu, v)| \leq k \sqrt{(B'u, u)} \sqrt{(B'v, v)} \leq k \sqrt{b'} \sqrt{(B'u, u)} \sqrt{(v, v)}$. Daraus für $v = Bu$ durch Quadrieren $(Bu, Bu) \leq k^2 b' (B'u, u)$, die verwendete Ungleichung.

Dann besitzt die Integralgleichung $\int_0^1 K(x, y; \varepsilon) \varphi(y) dy = \lambda \varphi(x)$ in einer festen Umgebung von $\varepsilon = 0$ ein reguläres Spektrum⁵⁾ mit dem festen Häufungspunkt Null⁶⁾.

Als ein Beispiel zu Satz 3 betrachten wir das Eigenwertproblem der Integralgleichung mit dem Kern $K(x, y; \varepsilon) = \frac{1}{2}(x+y) - \frac{1}{2}|x-y| - \frac{\varepsilon}{1+\varepsilon}xy$, $0 \leq x, y \leq 1$ aus der Einleitung. Die Voraussetzungen 1 und 2 von Satz 3 sind offenbar erfüllt. Um das Erfülltsein der Voraussetzung 3 (zweite Ungleichung) zu zeigen, beachte man

$$\begin{aligned} \int_0^1 \int_0^1 K(x, y; 0) u(x) \overline{u(y)} dx dy &= \int_0^1 u(x) \left| \int_0^x y \overline{u(y)} dy + x \int_x^1 \overline{u(y)} dy \right| dx \\ &= \int_0^1 \left| \int_x^1 u(y) dy \right|^2 dx. \end{aligned}$$

Es ist

$$\begin{aligned} &\left| \int_0^1 \int_0^1 K(x, y; \varepsilon) u(x) \overline{u(y)} dx dy \right| \\ &\leq \left| \int_0^1 \int_0^1 \frac{-\varepsilon}{1+\varepsilon} xy u(x) \overline{u(y)} dx dy \right| + \int_0^1 \int_0^1 K(x, y; 0) u(x) \overline{u(y)} dx dy. \end{aligned}$$

Für komplexes ε mit $|\varepsilon| \leq q < 1$ wird also

$$\begin{aligned} &\left| \int_0^1 \int_0^1 K(x, y; \varepsilon) u(x) \overline{u(y)} dx dy \right| \\ &\leq \frac{1}{1-q} \left| \int_0^1 x u(x) dx \right|^2 + \int_0^1 \int_0^1 K(x, y; 0) u(x) \overline{u(y)} dx dy. \end{aligned}$$

Wegen

$$\begin{aligned} \left| \int_0^1 x u(x) dx \right|^2 &= \left(\int_0^1 \left| \int_x^1 u(y) dy \right| dx \right)^2 \leq \int_0^1 \left| \int_x^1 u(y) dy \right|^2 dx \\ &= \int_0^1 \int_0^1 K(x, y; 0) u(x) \overline{u(y)} dx dy \end{aligned}$$

folgt

$$\begin{aligned} &\left| \int_0^1 \int_0^1 K(x, y; \varepsilon) u(x) \overline{u(y)} dx dy \right| \\ &\leq \left(\frac{1}{1-q} + 1 \right) \cdot \int_0^1 \int_0^1 K(x, y; 0) u(x) \overline{u(y)} dx dy. \end{aligned}$$

Es ist also auch Voraussetzung 3 erfüllt und daher hat der Integraloperator mit dem Kern $K(x, y; \varepsilon)$ ein reguläres Spektrum⁵⁾ mit dem festen Häufungspunkt Null.

§ 3.

Reguläre, nichtbeschränkte Operatoren mit festem Definitionsbereich

In diesem Paragraphen handelt es sich um nichtbeschränkte Operatoren $A(\varepsilon)$, die in einem festen, d. h. von ε unabhängigen Definitionsbereich \mathfrak{A} erklärt und symmetrisch sind. In Satz 5 der dritten Mitteilung haben wir ein hinreichendes Kriterium dafür kennengelernt, daß für einen solchen Operator $A(\varepsilon)u$ ein reguläres Element¹³⁾ darstellt, wenn u ein (von ε unabhängiges) Element aus \mathfrak{A} ist. Wir werden zunächst zeigen, daß dieses Kriterium auch notwendig ist (Hilfssatz 4) und dann in Hilfssatz 5 ein weiteres notwendiges und hinreichendes Kriterium herleiten.

Hilfssatz 4. *In einer Umgebung von $\varepsilon = 0$ sei $A(\varepsilon)$ in \mathfrak{A} ein symmetrischer Operator und $A(\varepsilon)u$ ein reguläres Element für jedes von ε unabhängige u aus \mathfrak{A} ; der Operator $A(0)$ sei in \mathfrak{A} selbstadjungiert. Dann gibt es in \mathfrak{A} symmetrische Operatoren $A_0 = A(0)$, A_1, A_2, \dots und eine Zahl k mit $|A_n u| \leq k^n \{|A_0 u| + |u|\}$; $n = 1, 2, \dots$ und $A(\varepsilon)u = A_0 u + \varepsilon A_1 u + \dots$ für alle u aus \mathfrak{A} .*

Beweis. Nach Voraussetzung ist $A(\varepsilon)u = u_0 + \varepsilon u_1 + \varepsilon^2 u_2 + \dots$. Also ist jedem u aus \mathfrak{A} (und jedem n) eindeutig ein u_n zugeordnet; zu zeigen ist, daß diese Zuordnung durch einen Operator $A_n u = u_n$ geschieht, d. h. daß sie linear ist. Aus $A(\varepsilon)v = v_0 + \varepsilon v_1 + \varepsilon^2 v_2 + \dots$ folgt aber, daß $\alpha u + \beta v$ dem Element $\alpha u + \beta v$ zugeordnet ist, also ist A_n ein Operator: $A(\varepsilon)u = A_0 u + \varepsilon A_1 u + \varepsilon^2 A_2 u + \dots$. Aus $(A(\varepsilon)u, v) = (u, A(\varepsilon)v)$ folgt $(A_n u, v) = (u, A_n v)$, also ist A_n symmetrisch. Sei $R = (A_0 + i)^{-1}$. Dann ist $A(\varepsilon)Rx = A_0 Rx + \varepsilon A_1 Rx + \dots$ für jedes x aus \mathfrak{H} . Für jedes ε aus einer Umgebung von $\varepsilon = 0$ ist $A(\varepsilon)R$ ein beschränkter Operator¹⁴⁾, und es ist $A(\varepsilon)Rx$ für jedes x aus \mathfrak{H} ein reguläres Element. Also ist¹⁵⁾ $|A_n Rx| \leq K^{n+1}|x|$ mit geeigneter Zahl K . Mit $Rx = u$, $x = A_0 u + iu$ gilt also für alle u aus \mathfrak{A} die Ungleichung $|A_n u| \leq K^{n+1}\{|A_0 u| + |u|\}$, die offenbar mit der Behauptung gleichwertig ist.

Die Tatsache, daß $A(\varepsilon)$ in einem festen Definitionsbereich selbstadjungiert ist und für jedes u dieses Bereiches $A(\varepsilon)u$ regulär ausfällt, kann auch an der Reziproken $R(\varepsilon)$ von $A(\varepsilon)$ (wenn eine solche existiert) erkannt werden. Es gilt nämlich

Hilfssatz 5. *Es sei $A(\varepsilon)$ in einer Umgebung von $\varepsilon = 0$ regulär und selbstadjungiert. Außerdem besitze $A_0 = A(0)$ eine beschränkte Reziproke R_0 , so daß¹⁵⁾ auch $A(\varepsilon)$ eine reguläre beschränkte Reziproke $R(\varepsilon) = R_0 + \varepsilon R_1 + \dots$ besitzt. Dann bestehen in einer Umgebung von $\varepsilon = 0$ Unabhängigkeit von ε*

¹³⁾ Definition 2 der ersten Mitteilung.

¹⁴⁾ Hilfssatz 2 der dritten Mitteilung.

¹⁵⁾ Satz 1 der dritten Mitteilung.

für den Definitionsbereich von $A(\varepsilon)$ und Regularität von $A(\varepsilon)u$ für jedes von ε unabhängige u aus diesem Definitionsbereich genau dann, wenn es eine Zahl C gibt, mit der $|R_n x| \leq C^n |R_0 x|$; $n = 1, 2, \dots$ für alle x aus \mathfrak{H} gilt.

Beweis. I. Die Bedingung $|R_n x| \leq C^n |R_0 x|$ ist notwendig. Dazu beachte man, daß $T(\varepsilon)x = A(\varepsilon)R_0 x$ für jedes Element x aus \mathfrak{H} erklärt ist und in einer Umgebung von $\varepsilon = 0$ ein reguläres Element darstellt. Nach Hilfssatz 2 der dritten Mitteilung ist $A(\varepsilon)R_0$ für jedes ε dieser Umgebung ein beschränkter Operator. Also ist (Definition 3' der ersten Mitteilung), $T(\varepsilon)$ ein regulärer beschränkter Operator. Da $T(0) = 1$ gewiß eine beschränkte Reziproke besitzt, so hat $T(\varepsilon)$ eine reguläre beschränkte Reziproke (Hilfssatz 4 der dritten Mitteilung): $S(\varepsilon) = S_0 + \varepsilon S_1 + \dots$. Es ist $S(\varepsilon)x = A_0 R(\varepsilon)x$. Für alle x aus \mathfrak{H} und alle u aus \mathfrak{U} gilt $(S(\varepsilon)x, u) = (R(\varepsilon)x, A_0 u)$ und daraus $(S_n x, u) = (R_n x, A_0 u)$. Da A_0 in \mathfrak{U} selbstadjungiert ist, liegt $R_n x$ in \mathfrak{U} und es ist $S_n = A_0 R_n$. Da $S(\varepsilon)$ regulär und beschränkt ist, gibt es⁷⁾ ein K mit $|S_n x| \leq K^{n+1} |x|$. Es ist $(u, S_n y) = (u, A_0 R_n y) = (R_n A_0 u, y)$ für alle u aus \mathfrak{U} und y aus \mathfrak{H} . Daher $|(R_n A_0 u, y)| \leq |u| \cdot K^{n+1} |y|$. Setzt man darin $y = R_n A_0 u$, so entsteht $|R_n A_0 u| \leq K^{n+1} |u|$ oder $|R_n x| \leq K^{n+1} |R_0 x|$ wie behauptet.

II. Die Bedingung $|R_n x| \leq C^n |R_0 x|$ ist hinreichend. Wir zeigen zuerst, daß $R_n x$ ein Element aus dem Definitionsbereich \mathfrak{U} von A_0 ist, wenn x ein beliebiges Element aus \mathfrak{H} ist. (Diese Tatsache ist übrigens in dem unten bewiesenen Hilfssatz 8 enthalten.) Aus $|R_n x| \leq C^n |R_0 x|$ folgt $|R_n A_0 u| \leq C^n |u|$ für alle u aus \mathfrak{U} . Der Operator $S_n^* = R_n A_0$, zunächst in \mathfrak{U} erklärt, kann zu einem in \mathfrak{H} erklärten, beschränkten Operator fortgesetzt werden. Seine (beschränkte) Adjungierte heiße S_n . Dann ist $(A_0 u, R_n x) = (u, S_n x)$ für alle u aus \mathfrak{U} und jedes x aus \mathfrak{H} . Da A_0 in \mathfrak{U} selbstadjungiert ist, folgt, daß $R_n x$ in \mathfrak{U} liegt und daß $S_n = A_0 R_n$ ist. Aus $|S_n x| \leq C^n |x|$ folgt sogar $|A_0 R_n x| \leq C^n |x|$. Es sei nun z ein beliebiges Element aus dem Definitionsbereich von $A(\varepsilon)$. Dann ist $z = (R_0 + \varepsilon R_1 + \dots)y$ mit geeignetem (von ε abhängigem) y aus \mathfrak{H} . Setzt man $z_n = (R_0 + \varepsilon R_1 + \dots + \varepsilon^n R_n)y$, so gilt $\lim_{n \rightarrow \infty} z_n = z$ und $A_0 z_n = y + \varepsilon A_0 R_1 y + \dots + \varepsilon^n A_0 R_n y$ besitzt (wegen $|A_0 R_n y| \leq C^n |y|$) auch einen Grenzwert, wenn n gegen ∞ rückt. Da A_0 in \mathfrak{U} abgeschlossen ist, liegt z in \mathfrak{U} . Also ist der Definitionsbereich von $A(\varepsilon)$ in \mathfrak{U} enthalten. Aber auch das Umgekehrte ist richtig. Sei nämlich u ein Element von \mathfrak{U} . Da $1 + \varepsilon A_0 R_1 + \varepsilon^2 A_0 R_2 + \dots$ ein regulärer beschränkter Operator ist, besitzt er (Hilfssatz 4 der dritten Mitteilung) eine (reguläre) beschränkte Reziproke $B(\varepsilon)$. Setzt man $B(\varepsilon)A_0 u = y$, so wird $A_0 u = (1 + \varepsilon A_0 R_1 + \varepsilon^2 A_0 R_2 + \dots)y$, $u = (R_0 + \varepsilon R_1 + \varepsilon^2 R_2 + \dots)y$, d. h. aber u liegt in dem Definitionsbereich von $A(\varepsilon)$. Der Definitionsbereich von $A(\varepsilon)$ fällt also mit \mathfrak{U} zusammen, er ist von ε unabhängig. — Wir haben noch

zu zeigen, daß $A(\varepsilon)u$ ein reguläres Element ist für jedes (von ε unabhängige) u aus \mathfrak{A} . Dazu bedenke man, daß $A(\varepsilon)R_0 = B(\varepsilon)$ ein regulärer beschränkter Operator ist. Also ist gewiß $A(\varepsilon)R_0x = B(\varepsilon)x$ für jedes x aus \mathfrak{S} ein reguläres Element. Also auch $A(\varepsilon)u$ für jedes u aus \mathfrak{A} .

Wir sind jetzt in der Lage, das hauptsächliche Ziel dieses Paragraphen zu beweisen, nämlich

Satz 4. *Für eine Umgebung von $\varepsilon = 0$ sei $A(\varepsilon)$ in einem festen, von ε unabhängigen Definitionsbereich \mathfrak{A} symmetrisch. $A(\varepsilon)u$ sei regulär für jedes u aus \mathfrak{A} . Der ungestörte Operator $A(0)$ sei in \mathfrak{A} selbstadjungiert und besitze ein diskretes Spektrum. Dann hat $A(\varepsilon)$ in einer Umgebung von $\varepsilon = 0$ ein reguläres diskretes Spektrum (Definition 1).*

Beweis. Nach Satz 4 der dritten Mitteilung ist $A(\varepsilon)$ für alle ε einer Umgebung von $\varepsilon = 0$ selbstadjungiert und regulär. Da $A(0)$ ein diskretes Spektrum hat, gibt es eine reelle Zahl d , so daß $A(0) + d$ eine beschränkte Reziproke R_0 besitzt. Dann hat $A(\varepsilon) + d$ in einer Umgebung von $\varepsilon = 0$ eine reguläre beschränkte und symmetrische Reziproke $R(\varepsilon) = R_0 + \varepsilon R_1 + \dots$ und es ist R_0 vollstetig und symmetrisch. Nach Hilfssatz 5 gibt es eine Konstante C , so daß $|R_n x| \leq C^n |R_0 x|$ gilt. Nach Satz 1 besitzt also $R(\varepsilon)$ im Intervall $-\frac{1}{2C} < \varepsilon < \frac{1}{2C}$ ein reguläres Spektrum mit dem festen Häufungspunkt Null. Da R_0 nicht den Eigenwert Null hat, sind alle Eigenwerte von $R(\varepsilon)$ überall (in dieser Umgebung von $\varepsilon = 0$) von Null verschieden. Also hat $A(\varepsilon) + d$ und damit auch $A(\varepsilon)$ in $-\frac{1}{2C} < \varepsilon < \frac{1}{2C}$ ein reguläres diskretes Spektrum.

Für halbbeschränkte Operatoren haben wir früher (Satz 6 der dritten und Satz 3 der vierten¹⁸⁾ Mitteilung) Regularitätskriterien auch in solchen Fällen gewonnen, wo der Definitionsbereich vom Störungsparameter ε abhängt; erst der Definitionsbereich der zugehörigen Form war von ε unabhängig. Auch in diesen Fällen kann man, falls der ungestörte Operator ein diskretes Spektrum besitzt, für den gestörten Operator ein reguläres diskretes Spektrum nachweisen. Wir zeigen das für Satz 3 der vierten Mitteilung als dem allgemeineren Satz und beweisen

Satz 5. *Fügt man den Voraussetzungen von Satz 3 der vierten Mitteilung die Voraussetzung hinzu, A_0 habe ein diskretes Spektrum, dann besitzt der in diesem Satz genannte Operator $A(\varepsilon)$ in einer Umgebung von $\varepsilon = 0$ ein reguläres diskretes Spektrum.*

Beweis. Mit den Bezeichnungen von Satz 3 der vierten Mitteilung [vgl. Formel (40), S. 372] wird $\left(A(\varepsilon) + \frac{a+1}{b}\right)^{-1} = R(\varepsilon) = R_0 + \varepsilon R_1 + \dots$

¹⁸⁾ Math. Annalen 117 (1940), S. 356–382.

ein regulärer, beschränkter und symmetrischer Operator und es ist $R(\varepsilon) = B(\varepsilon)R$, $R = (1 + a + bA_0)^{-1}$. Dabei gilt für $B(\varepsilon) = \sum_{r=0}^{\infty} \varepsilon^r B_r$, da es in \mathfrak{G} ein G -regulärer, G -beschränkter Operator ist, $[B, u, B, u] \leq K^2 [u, u]$ für jedes u aus \mathfrak{G} und daher $|[B, u, u]| \leq K' [u, u]$ oder $|[B, Rx, Rx]| \leq K' [Rx, Rx]$, d. h. $|[B, Rx, x]| \leq K' (Rx, x)$ oder $|(R, x, x)| \leq K' (Rx, x)$, $r = 1, 2, \dots$. Es ist $R_0 = b(bA_0 + a + 1)^{-1} = bR$, also $|(R, x, x)| \leq \frac{K'}{b} (R_0 x, x)$ oder $|(R, x, x)| \leq k' (R_0 x, x)$ mit geeignetem k . Da R_0 vollstetig ist, sind alle Voraussetzungen von Satz 2 erfüllt, $R(\varepsilon)$ besitzt also in einer Umgebung von $\varepsilon = 0$ ein reguläres Spektrum mit dem festen Häufungspunkt Null. Da R_0 nicht den Eigenwert Null hat, sind die Eigenwerte von $R(\varepsilon)$ überall (in einer Umgebung von $\varepsilon = 0$) von Null verschieden und es hat daher $R(\varepsilon)^{-1} = A(\varepsilon) + \frac{a+1}{b}$ und daher auch $A(\varepsilon)$ ein reguläres diskretes Spektrum.

Die Sätze 4 und 5 lassen erkennen, daß alle Typen von Differentialoperatoren, die wir in § 6 der vierten Mitteilung als regulär nachgewiesen haben, ein reguläres diskretes Spektrum besitzen, sofern sie im ungestörten Zustand ein diskretes Spektrum haben. Insbesondere gilt das natürlich für das in der Einleitung genannte Eigenwertproblem $u'' + \lambda u = 0$; $u(0) = 0$, $u'(1) + \varepsilon u(1) = 0$.

§ 4.

Sätze über Operatoren mit diskretem Spektrum und über vollstetige Operatoren.

Hilfssatz 6. *Ein in \mathfrak{A} selbstadjungierter Operator A hat dann und nur dann ein diskretes Spektrum, wenn aus jeder unendlichen Menge $\{u\}$ von Elementen aus \mathfrak{A} mit $|Au| + |u| \leq C$ eine konvergente Teilfolge ausgewählt werden kann.*

Beweis. Sei zunächst das Spektrum von A diskret und a eine reelle Zahl derart, daß $A + a$ eine beschränkte symmetrische Reziproke R hat. Für $x = Au + au$ und u aus $\{u\}$ wird $|x| \leq |Au| + |a||u| \leq (1 + |a|)C$. Da R vollstetig ist, kann man aus der Menge der x eine Teilfolge $x_n = Au_n + au_n$ auswählen, für die Rx_n konvergiert. Es ist also $Rx_n = u_n$ eine konvergente Teilfolge von $\{u\}$. — Wäre umgekehrt das Spektrum von A nicht diskret, dann gäbe es eine Zahl p , so daß der zu $-p < \lambda < p$ gehörige Spektralraum von A unendliche Dimension hätte. Es sei $\varphi_1, \varphi_2, \dots$ ein unendliches System orthogonaler normierter Elemente aus diesem Raum. Es wäre dann $|\varphi_n| = 1$, $|A\varphi_n| \leq p|\varphi_n|$, also $|A\varphi_n| + |\varphi_n| \leq 1 + p$. Man müßte aus der Menge der $\varphi_1, \varphi_2, \dots$ eine konvergente Teilfolge auswählen können, was unmöglich ist.

Hilfssatz 7. Es sei A in \mathfrak{A} selbstadjungiert und B in \mathfrak{A} symmetrisch; dann gibt es eine Zahl k , mit der $|Bu| \leq k\{|Au| + |u|\}$ gilt für alle u aus \mathfrak{A} .

Zum Beweis bilde man $R = (A + i)^{-1}$. Nach Hilfssatz 2 der dritten Mitteilung ist BR ein beschränkter Operator, $|BRx| \leq k|x|$. Mit $Rx = u$ wird $x = Au + iu$, also $|Bu| \leq k\{|Au| + |u|\}$ wie behauptet.

Hilfssatz 8. Es seien R und S zwei beschränkte, symmetrische Operatoren. Dann gilt $|Sx| \leq k|Rx|$ für alle x aus \mathfrak{H} genau dann, wenn der Wertebereich von S in dem von R enthalten ist.

Beweis. I. Wir führen den Beweis zunächst unter der zusätzlichen Annahme, daß aus $Rx = 0$ folge $x = 0$. Sei dann $|Sx| \leq k|Rx|$ und bezeichne A die (selbstadjungierte) Reziproke von R . Ihr Definitionsbereich \mathfrak{A} ist der Wertebereich von R . Aus $|SAu| \leq k|u|$ für alle u aus \mathfrak{A} folgt, daß der zunächst nur in \mathfrak{A} erklärte Operator $T = SA$ zu einem in \mathfrak{H} beschränkten Operator fortgesetzt werden kann. Sein Adjungierter heiße T^* . Dann ist $(SAu, x) = (u, T^*x)$ für alle u aus \mathfrak{A} , x aus \mathfrak{H} . Oder $(Au, Sx) = (u, T^*x)$ für alle u aus \mathfrak{A} . Da A selbstadjungiert ist, muß Sx in \mathfrak{A} liegen, also ist die Bedingung $|Sx| \leq k|Rx|$ hinreichend dafür, daß der Wertebereich von S in dem von R enthalten ist. — Die Bedingung ist aber auch notwendig. Denn wenn der Wertebereich von S in dem von R enthalten ist, dann ist AS ein in \mathfrak{H} erklärter, beschränkter¹⁴⁾ Operator: $|ASx| \leq k|x|$. Nun ist $(SAu, y) = (u, ASy)$ für u aus \mathfrak{A} , y aus \mathfrak{H} , also $|(SAu, y)| \leq k|u||y|$ und das wird mit $y = SAu$ soviel wie $|SAu| \leq k|u|$. Also $|Sx| \leq k|Rx|$ für alle x aus \mathfrak{H} .

II. Wir befreien uns von der Annahme $Rx \neq 0$ für $x \neq 0$. Und zwar sei zunächst $|Sx| \leq k|Rx|$. Wir bezeichnen mit \mathfrak{R} die Gesamtheit der x , für die $Rx = 0$ ist und mit $\tilde{\mathfrak{H}}$ den (abgeschlossenen) Raum aller \tilde{x} , die auf \mathfrak{R} senkrecht stehen; wir nehmen an, daß $\tilde{\mathfrak{H}}$ nicht nur das Nullelement enthält, weil sonst $R = 0$ und nichts zu beweisen wäre. Für jedes x aus \mathfrak{R} ist auch $Sx = 0$. Sowohl R als auch S bilden jedes Element aus $\tilde{\mathfrak{H}}$ auf ein Element von $\tilde{\mathfrak{H}}$ ab, sie können als beschränkte symmetrische Operatoren \tilde{R} und \tilde{S} in $\tilde{\mathfrak{H}}$ aufgefaßt werden, für die $|\tilde{S}\tilde{x}| \leq k|\tilde{R}\tilde{x}|$ gilt. Außerdem $\tilde{x} = 0$, wenn $R\tilde{x} = 0$. Nach dem in Schritt I Bewiesenen ist der Wertebereich von \tilde{S} in dem von \tilde{R} enthalten. Da die Wertebereiche von \tilde{S} und S bzw. von \tilde{R} und R gleich sind, ist auch gezeigt, daß der Wertebereich von S in dem von R enthalten ist. — Nehmen wir umgekehrt an, der Wertebereich von S sei in dem von R enthalten. Für jedes y aus \mathfrak{H} ist dann $Sy = Rz$ mit geeignetem z , also orthogonal zu jedem x aus \mathfrak{R} , d. h. aus $Rx = 0$ folgt wieder $Sx = 0$. Erklärt man \tilde{R} und \tilde{S} wie oben, dann ist der Wertebereich von \tilde{S} in dem von \tilde{R} enthalten, außerdem $\tilde{R}\tilde{x} \neq 0$ für $\tilde{x} \neq 0$. Also nach Schritt I zunächst $|\tilde{S}\tilde{x}| \leq k|\tilde{R}\tilde{x}|$ für alle \tilde{x} aus $\tilde{\mathfrak{H}}$ und daraus $|Sx| \leq k|Rx|$ für alle x aus \mathfrak{H} . Damit ist der Hilfssatz bewiesen.

Satz 6. Wenn es überhaupt einen in \mathfrak{A} selbstadjungierten Operator A mit diskretem Spektrum gibt, dann hat jeder in \mathfrak{A} oder in einem Teilraum von \mathfrak{A} selbstadjungierte Operator ein diskretes Spektrum.

Beweis. Es sei A in \mathfrak{A} selbstadjungiert, mit diskretem Spektrum, und es sei B ein in einem Teilraum \mathfrak{B} von \mathfrak{A} selbstadjungierter Operator. Um zu zeigen, daß B ein diskretes Spektrum hat, muß man nach Hilfssatz 6 zeigen, daß jede unendliche Menge $\{u\}$ aus \mathfrak{B} mit $|Bu| + |u| \leq C$ eine konvergente Teilfolge u_n enthält. Nun ist nach Hilfssatz 7 jedenfalls $|Au| \leq k\{|Bu| + |u|\}$ für alle u aus \mathfrak{B} , also $|Au| \leq kC$, $|Au| + |u| \leq kC + C$ für alle u aus $\{u\}$. Da A ein diskretes Spektrum besitzt, läßt sich (Hilfssatz 6) aus $\{u\}$ eine konvergente Teilfolge auswählen, also hat (wieder Hilfssatz 6) auch B ein diskretes Spektrum.

Der entsprechende Satz über vollstetige Operatoren lautet so:

Satz 7. Es sei R ein beschränkter, symmetrischer Operator mit dem Wertebereich \mathfrak{B} . Wenn es dann überhaupt einen vollstetigen symmetrischen Operator gibt, dessen Wertebereich den Bereich \mathfrak{B} enthält, dann ist auch R vollstetig.

Beweis. Es sei V ein vollstetiger (also auch beschränkter) symmetrischer Operator mit einem \mathfrak{B} enthaltenden Wertebereich. Dann gibt es nach Hilfssatz 8 eine Konstante k , mit der $|Rx| \leq k|Vx|$ gilt. Also ist auch R vollstetig.

§ 5.

Gegenbeispiele.

1. Wenn ein Operator $A(\varepsilon)$ für eine Umgebung⁸⁾ von $\varepsilon = 0$ in einem von ε unabhängigen Definitionsbereich selbstadjungiert ist und wenn $A(\varepsilon)u$ für jedes u aus \mathfrak{A} ein reguläres Element darstellt, dann ist $A(\varepsilon)$ in einer Umgebung von $\varepsilon = 0$ ein regulärer Operator. Das wurde in Satz 4 der dritten Mitteilung bewiesen; es wurde dort sogar gezeigt, daß es genügt, die Selbstadjungiertheit von $A(0)$ in \mathfrak{A} vorauszusetzen, die von $A(\varepsilon)$ kann gefolgert werden. Umgekehrt aber kann nicht behauptet werden, daß für einen Operator $A(\varepsilon)$, regulär und selbstadjungiert in einem festen Teilraum \mathfrak{A} , auch immer $A(\varepsilon)u$ ein reguläres Element sei, wenn u ein beliebiges von ε unabhängiges Element aus \mathfrak{A} ist. Dafür das folgende Gegenbeispiel: Es sei \mathfrak{H} der reelle Hilbertsche Raum aller reellen $x(t)$ mit $\int_1^\infty x^2(t) dt < \infty$; $(x, y) = \int_1^\infty x(t)y(t) dt$. Es sei der Operator $A(\varepsilon)$ für $-1 < \varepsilon < 1$ erklärt durch $A(\varepsilon)x = \frac{2t^{1+\varepsilon}}{1+t^2} x(t)$. Er ist selbstadjungiert im Bereich $\mathfrak{A}(\varepsilon)$ aller x aus \mathfrak{H} , für die $\int_1^\infty \left(\frac{t^{1+\varepsilon}}{1+t^2}\right)^2 x^2(t) dt < \infty$ ist. Dieses

$\mathfrak{H}(\varepsilon)$ hängt aber gar nicht von ε ab, sondern ist identisch mit $\mathfrak{H} = \mathfrak{H}(0)$, d. h. mit der Gesamtheit aller x aus \mathfrak{H} , für die $\int_1^\infty t^2 x^2(t) dt < \infty$ gilt. Es ist also $A(\varepsilon)$ in \mathfrak{H} selbstadjungiert. Um einzusehen, daß es auch regulär ist in $-1 < \varepsilon < 1$, brauchen wir nur zu zeigen, daß die Reziproke $R(\varepsilon)x = \frac{1+t^2}{2t^2} x(t)$ ein regulärer beschränkter Operator ist. Für jedes ε und für alle $t \geq 1$ ist $\frac{1+t^2}{2t^{1+\varepsilon}} = \frac{1}{2t} + \frac{1}{2t} e^{-\varepsilon^2 \log t} = \frac{1}{t} + \frac{1}{2} \sum_{\nu=1}^{\infty} (-1)^\nu \frac{\log^\nu t}{\nu! t} \varepsilon^{2\nu}$.

Wir setzen $R_0 x = \frac{1}{t} x(t)$, $R_\mu x = \frac{1}{2} \frac{(-1)^\nu \log^\nu t}{\nu! t} x(t)$, wenn $\mu = 2\nu$ und $R_\mu = 0$, wenn $\mu = 2\nu - 1$ ($\nu = 1, 2, \dots$). Wir wollen zeigen, daß $S(\varepsilon) = R_0 + \varepsilon R_1 + \varepsilon^2 R_2 + \dots$ in $-1 < \varepsilon < 1$ ein regulärer beschränkter und symmetrischer Operator ist, der mit $R(\varepsilon)$ übereinstimmt. Es ist $\max_{1 \leq t} \frac{1}{2} \frac{\log^\nu t}{\nu! t} = \frac{1}{2} \frac{\nu^\nu e^{-\nu}}{\nu!} < C$, woraus folgt¹⁷⁾ $|R_\mu x| \leq C|x|$. Also ist $S(\varepsilon)$ in $-1 < \varepsilon < 1$ regulär, beschränkt (und natürlich symmetrisch). Um $S(\varepsilon) = R(\varepsilon)$ zu zeigen, bilden wir $|(R(\varepsilon) - S(\varepsilon)x)| = \lim_{n \rightarrow \infty} d_n$ mit

$$\begin{aligned} d_n &= \left| (R(\varepsilon) - \sum_{\nu=0}^n R_\nu \varepsilon^\nu) x \right| = \left[\int_1^\infty \left(\frac{1+t^2}{2t^{1+\varepsilon}} - \frac{1}{t} - \frac{1}{2} \sum_{\nu=1}^n (-1)^\nu \frac{\log^\nu t}{\nu! t} \varepsilon^{2\nu} \right)^2 x^2(t) dt \right]^{\frac{1}{2}} \\ &= \left[\int_1^\infty \left(\frac{1}{2} \sum_{\nu=n+1}^{\infty} (-1)^\nu \frac{\log^\nu t}{\nu! t} \varepsilon^{2\nu} \right)^2 x^2(t) dt \right]^{\frac{1}{2}} \leq \sum_{\nu=n+1}^{\infty} \left[\int_1^\infty \left(\frac{1}{2} \frac{\log^\nu t}{\nu! t} \varepsilon^{2\nu} x(t) \right)^2 dt \right]^{\frac{1}{2}} \\ &\leq C \sum_{\nu=n+1}^{\infty} \varepsilon^{2\nu} \left[\int_1^\infty x^2(t) dt \right]^{\frac{1}{2}} = |x| \cdot C \cdot \sum_{\nu=n+1}^{\infty} \varepsilon^{2\nu}. \end{aligned}$$

Also $\lim_{n \rightarrow \infty} d_n = 0$, $R(\varepsilon) = S(\varepsilon)$.

Damit ist gezeigt, daß $A(\varepsilon)$ in \mathfrak{H} regulär und selbstadjungiert ist. Bleibt noch zu zeigen, daß nicht für jedes (von ε unabhängige) u aus \mathfrak{H} das Element $A(\varepsilon)u$ ein reguläres Element ist. Dazu sei $u = u(t)$ ein Element aus \mathfrak{H} , also $\int_1^\infty t^2 u^2(t) dt < \infty$, für das $\int_1^\infty \log^2 t \cdot t^2 u^2(t) dt = \infty$ ausfällt. Wäre $A(\varepsilon)u$ regulär, $A(\varepsilon)u = u_0 + \varepsilon u_1 + \varepsilon^2 u_2 + \dots$ in einer Umgebung von $\varepsilon = 0$, dann müßte dort $u = (R_0 + \varepsilon R_1 + \varepsilon^2 R_2 + \dots)(u_0 + \varepsilon u_1 + \varepsilon^2 u_2 + \dots)$ sein. Wegen $R_1 = 0$ wäre $u = R_0 u_0$, also $u_0 = tu(t)$ und $0 = R_2 u_0 + R_0 u_2$, also $R_2 u_0 = -R_0 u_2$, $R_2 u_0 = -\frac{1}{2} \frac{\log t}{t} tu(t) = -\frac{1}{2} \log t \cdot u(t)$. Wegen $\int_1^\infty t^2 \log^2 t \times$

¹⁷⁾ Die Absolutstriche sind im Sinne des Hilbertschen Raumes gemeint; z. B. $|x| = \left[\int_1^\infty x^2(t) dt \right]^{\frac{1}{2}}$.

$\times u^2(t)dt = \infty$ liegt $R_2 u_0$ nicht in \mathfrak{H} . Wegen $R_2 u_0 = -R_0 u_2$ liegt es in \mathfrak{H} , Widerspruch. Damit ist das Gegenbeispiel gewonnen. Man bemerkt übrigens, daß es wesentlich an der Beschränkung auf *reelle* ε hängt: in einer noch so kleinen aber *komplexen* Umgebung von $\varepsilon = 0$ wäre $A(\varepsilon)$ nicht in einem festen Definitionsbereich regulär.

Der Operator $R(\varepsilon) = R_0 + \varepsilon R_1 + \varepsilon^2 R_2 + \dots$ ist zugleich ein Beispiel für einen (in einer Umgebung⁸⁾ von $\varepsilon = 0$) regulären beschränkten und symmetrischen Operator, dessen Wertebereich ein fester von ε unabhängiger Teilraum \mathfrak{H} ist, während $R_1 x$ nicht für alle x in \mathfrak{H} liegt. Denn setzt man $x = tu(t)$, so wird $R_2 x = -\frac{1}{2} \log t \cdot u(t)$ und dieses Element liegt *nicht* in \mathfrak{H} .

2. Unter den Voraussetzungen des Satzes 1 kann zwar, wie eben dieser Satz lehrt, die Gesamtheit der Eigenwerte $\lambda_1(\varepsilon), \lambda_2(\varepsilon), \dots$ über ein gemeinsames Intervall regulär analytisch fortgesetzt werden. Aber es braucht kein (noch so kleines) *gemeinsames* Intervall $|\varepsilon| < p$ zu geben, in dem die Potenzreihenentwicklung (um $\varepsilon = 0$) für jedes der $\lambda_1(\varepsilon), \lambda_2(\varepsilon), \dots$ konvergent wäre. Das erkennt man an folgendem Beispiel: Wir legen den reellen Hilbertschen

Raum \mathfrak{H}_0 aller reellen Vektoren $x = (x_1, x_2, \dots)$ mit $\sum_{r=1}^{\infty} x_r^2 < \infty$ zugrunde. Es seien R_0, R_1, R_2, \dots Matrizen, und zwar sei $R_2 = R_3 = \dots = 0$ und

$$R_0 = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \lambda_3 & \\ & & & \lambda_4 \ddots \end{pmatrix}, \quad R_1 = \begin{pmatrix} 0 & \mu_1 & & \\ \mu_1 & 0 & & \\ & & 0 & \mu_2 \\ & & 0 & \mu_2 & 0 \ddots \end{pmatrix}.$$

Dabei werden wir $\lambda_r = \mu_r = \frac{1}{r}$ wählen. Offenbar stellen R_0 und R_1 in \mathfrak{H}_0 symmetrische Operatoren dar und es ist R_0 vollstetig. Aus $R_0 x = (\lambda_1 x_1, \lambda_2 x_2, \dots)$ und $R_1 x = (\mu_1 x_2, \mu_1 x_1, \dots, \mu_n x_{2n}, \mu_n x_{2n-1}, \dots)$

folgt $(R_0 x, R_0 x) = \sum_{r=1}^{\infty} \lambda_r^2 x_r^2$, $(R_1 x, R_1 x) = \sum_{r=1}^{\infty} \mu_{\frac{r+1}{2}}^2 x_r^2$, also $(R_0 x, R_0 x) = \sum_{r=1}^{\infty} \frac{x_r^2}{r^2}$, $(R_1 x, R_1 x) = \sum_{r=1}^{\infty} \frac{x_r^2}{\left[\frac{r+1}{2}\right]^2} \leq x_1^2 + \sum_{r=2}^{\infty} \frac{x_r^2}{\left(\frac{r+1}{2} - 1\right)^2} \leq x_1^2 + \sum_{r=2}^{\infty} \frac{16 x_r^2}{r^2} \leq 16 (R_0 x, R_0 x)$. Also $|R_1 x| \leq 4 |R_0 x|$, man kann $k = 4$ in

Satz 1 setzen und alle Voraussetzungen dieses Satzes sind erfüllt. Die Eigenwerte $\lambda_n(\varepsilon)$ von $R_0 + \varepsilon R_1$ können angegeben werden. Es sind dies offenbar die Wurzeln λ der Gleichungen $\begin{vmatrix} \lambda_{2n-1} - \lambda & \varepsilon \mu_n \\ \varepsilon \mu_n & \lambda_{2n} - \lambda \end{vmatrix} = 0$; $n = 1, 2, \dots$

Es wird

$$\lambda_{2n-1}(\varepsilon) = \frac{\lambda_{2n-1} + \lambda_{2n}}{2} + \frac{\lambda_{2n-1} - \lambda_{2n}}{2} \left[1 + \varepsilon^2 \left(\frac{2\mu_n}{\lambda_{2n-1} - \lambda_{2n}} \right)^2 \right]^{\frac{1}{2}},$$

$$\lambda_{2n}(\varepsilon) = \frac{\lambda_{2n-1} + \lambda_{2n}}{2} - \frac{\lambda_{2n-1} - \lambda_{2n}}{2} \left[1 + \varepsilon^2 \left(\frac{2\mu_n}{\lambda_{2n-1} - \lambda_{2n}} \right)^2 \right]^{\frac{1}{2}}.$$

Es haben $\lambda_{2n-1}(\varepsilon)$ und $\lambda_{2n}(\varepsilon)$ um $\varepsilon = 0$ die (genaue) Konvergenzstrecke

$$\frac{\lambda_{2n-1} - \lambda_{2n}}{2\mu_n} = \frac{\frac{1}{2n-1} - \frac{1}{2n}}{\frac{2}{n}} = \frac{1}{2} \left(\frac{n}{2n-1} - \frac{1}{2} \right), \text{ die für } n \rightarrow \infty \text{ gegen Null}$$

rückt. Damit ist das Gegenbeispiel angegeben.

(Eingegangen am 2. 5. 1942.)

Integralsätze über Polynome mit lauter reellen Nullstellen.

Von

Stephan Lipka in Szeged (Ungarn).

Wir werden in dieser Arbeit Integralsätze ableiten, die sich auf Polynome mit lauter reellen Nullstellen beziehen, und über die Verteilung der Nullstellen der Ableitungen des Polynoms einen Aufschluß geben.

§ 1.

Satz 1. Es bedeute $f(x)$ ein Polynom n -ten Grades mit lauter reellen Nullstellen: $x_1 \leq x_2 \leq \dots \leq x_n$. Die Nullstellen der Ableitung $f'(x)$ seien: $y_1 \leq y_2 \leq \dots \leq y_{n-1}$. Ist nun

$$x_2 - x_1 \geq x_3 - x_2$$

und ist $f(x) > 0$ im Intervall (x_1, x_2) , so gilt die Integralungleichung

$$(1) \quad \int_{y_1}^{y_2} f(x) dx \geq 0.$$

(Das Gleichheitszeichen gilt nur im Falle: $n = 3$, $x_2 - x_1 = x_3 - x_2$; $x_1 = x_2 = x_3$.)

Beweis. Den Satz beweisen wir erstens für $n = 3$. Man kann ohne Beschränkung der Allgemeinheit voraussetzen, daß $x_1 = 0$, $x_2 = 1$ ist. Dann wird

$$f(x) = x(x-1)(x-\alpha),$$

wobei nach Voraussetzung $1 \leq \alpha \leq 2$ ist,

$$(2) \quad y_1 = \frac{1+\alpha-\sqrt{\Delta}}{3}, \quad y_2 = \frac{1+\alpha+\sqrt{\Delta}}{3}, \quad \Delta = (\alpha+1)^2 - 3\alpha > 0;$$

und das Integral (1) hat den folgenden Wert:

$$\int_{y_1}^{y_2} (x^3 - (1+\alpha)x^2 + \alpha x) dx = \frac{1}{4}(y_2^4 - y_1^4) - \frac{1}{3}(y_2^3 - y_1^3)(1+\alpha) + \frac{\alpha}{2}(y_2^2 - y_1^2).$$

Wendet man nun die Formel $y_2^k - y_1^k = (y_2 - y_1) \sum_{i=1}^k y_2^{k-i} y_1^{i-1}$ für die rechte Seite der vorigen Gleichung an, so folgt leicht nach (2)

$$\int_{y_1}^{y_2} f(x) dx = \frac{2\sqrt{\Delta}}{81}(\alpha+1)(\alpha-2)(-2\alpha+1).$$

Da nach Voraussetzung $1 \leq \alpha \leq 2$ gilt, so ergibt sich unmittelbar aus der letzteren Formel die Richtigkeit des Satzes für $n = 3$.

Es sei nun $n > 3$ und man setze

$$f(x) = \varphi(x) \psi(x),$$

wobei

$$\varphi(x) = (x - x_1)(x - x_2)(x - x_3)$$

und

$$\psi(x) = \prod_{i=4}^n (x - x_i)$$

ist. Bedeuten η_1, η_2 die Wurzeln von $\varphi'(x)$, so gelten die Ungleichungen

$$(3) \quad \eta_1 > y_1, \quad \eta_2 > y_2,$$

wobei y_1, y_2 die Wurzeln von $f'(x)$ sind. Betrachtet man nämlich den Wert von $\frac{f'(x)}{f(x)}$ für $x = \eta_1$, so folgt, da $\varphi'(\eta_1) = 0$ und $\eta_1 < x_i$ für $i > 3$ ist, die Ungleichung

$$\frac{f'(\eta_1)}{f(\eta_1)} = \frac{\varphi'(\eta_1)}{\varphi(\eta_1)} + \frac{\psi'(\eta_1)}{\psi(\eta_1)} = \frac{\psi'(\eta_1)}{\psi(\eta_1)} = \sum_{i=4}^n \frac{1}{\eta_1 - x_i} < 0.$$

Für $x_1 < x < y_1$ ist aber $\frac{f'(x)}{f(x)} > 0$, also muß $\eta_1 > y_1$ sein. Ebenso beweist man die Richtigkeit der zweiten Ungleichung von (3). Da der Satz für $n = 3$ richtig ist, so besteht

$$\int_{\eta_1}^{y_1} \varphi(x) dx \geq 0.$$

Da $\varphi(x)$ in (x_1, x_2) und (x_2, x_3) positiv bzw. negativ ist, so folgt nach (3)

$$(4) \quad \int_{y_1}^{y_2} \varphi(x) dx > \int_{\eta_1}^{y_1} \varphi(x) dx.$$

Das Polynom $\psi(x) = \prod_{i=4}^n (x - x_i)$ ist im Intervall (y_1, x_2) positiv und erreicht wegen $x_2 \leq x_4 \leq x_5 \leq \dots \leq x_n$ für $x = x_2$ sein Minimum in (y_1, x_2) . Hieraus folgt unmittelbar die Ungleichung

$$(5) \quad \int_{y_1}^{x_2} \varphi(x) \psi(x) dx > \psi(x_2) \int_{y_1}^{x_2} \varphi(x) dx.$$

Da $\varphi(x)$ und $\psi(x)$ im Intervall (x_2, y_2) negativ bzw. positiv sind, ferner $\psi(x)$ in (x_2, y_2) sein Maximum für $x = x_2$ erreicht, so folgt die Ungleichung

$$(6) \quad \int_{x_2}^{y_2} \varphi(x) \psi(x) dx > \psi(x_2) \int_{x_2}^{y_2} \varphi(x) dx.$$

Addiert man die entsprechenden Seiten von (5) und (6), so gewinnt man die Ungleichung:

$$\int_{y_1}^{y_2} \varphi(x) \psi(x) dx > \varphi(x_2) \int_{y_1}^{y_2} \varphi(x) dx.$$

Hieraus folgt nach (4) die Ungleichung

$$\int_{y_1}^{y_2} \varphi(x) \psi(x) dx > 0,$$

w. z. b. w.

§ 2.

Der Grad von $f(x)$ sei nun $n > 3$. Die Wurzeln der zweiten Ableitung $f''(x)$ bezeichnen: $z_1 \leq z_2 \leq \dots \leq z_n$. Sind die drei ersten Wurzeln von $f(x)$ voneinander verschieden, d. h. $x_1 < x_2 < x_3$, so gilt der

Satz 2. *Genügen die Wurzeln von $f'(x)$ der Bedingung*

$$y_2 - y_1 \geq y_3 - y_2$$

und liegt die Wurzel z_1 im Intervall (x_2, x_3) , so liegt die Wurzel z_2 auch im Intervall (x_2, x_3) . Liegt z_2 außerhalb des Intervalles (x_2, x_3) , so liegt z_1 auch außerhalb von (x_2, x_3) .

Beweis. Wir betrachten das Integral

$$\int_{z_1}^{z_2} f'(x) dx.$$

Da $f(x)$ im Intervall (x_1, x_2) positiv ist, so folgt, daß $f'(x)$ in (y_1, y_2) negativ ist. Hieraus folgt nach dem Satz 1, daß

$$(7) \quad \int_{z_1}^{z_2} f'(x) dx = f(z_2) - f(z_1) \leq 0$$

gilt. Liegt nun z_1 im Intervall (x_2, x_3) , so ist $f(z_1) < 0$ und daher folgt aus (7) die Ungleichung $f(z_2) < 0$. Die Wurzel z_2 liegt also auch in (x_2, x_3) . Liegt z_2 außerhalb des Intervalles (x_2, x_3) , so ist $f(z_2) > 0$ und nach (7) $f(z_1) > 0$, d. h. z_1 liegt außerhalb von (x_2, x_3) .

§ 3.

Hilfssatz I. *Es seien $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_k$ beliebige positive Zahlen. Man bilde das Polynom*

$$g(x) = (x + \alpha_1)(x + \alpha_2) \dots (x + \alpha_k).$$

Bezeichnet γ eine feste positive Zahl, für welche

$$1 \leq \gamma$$

ist, und betrachtet man alle Polynome $g(x)$, die den Bedingungen

$$(8) \quad \frac{1}{\alpha_1} + \frac{1}{\alpha_2} + \dots + \frac{1}{\alpha_k} = \gamma, \quad \alpha_2 \geq \beta > 0$$

genügen, so erreicht das Integral

$$\int_{-\beta}^2 g(x) (x-1) dx$$

sein Minimum für $\alpha_1 = \alpha_2 = \dots = \alpha_k = \frac{k}{\gamma}$, wenn $0 < \beta \leq \frac{k}{\gamma}$ ist.

Zum Beweis des Hilfssatzes wenden wir Induktion an, daher setzen wir voraus, daß der Satz richtig ist, wenn der Grad von $g(x) < k$ ist. Sind die Werte α_i nicht alle gleich, so muß nach (8)

$$\alpha_1 < \frac{k}{\gamma} < \alpha_k$$

sein. Wir wollen zeigen, daß in diesem Falle ein Polynom $g^*(x)$ existiert, das (8) genügt und für welches

$$\int_{-\beta}^2 g^*(x) (x-1) dx < \int_{-\beta}^2 g(x) (x-1) dx$$

gilt. Damit wird aber der Satz bewiesen. Man bilde das harmonische Mittel:

$$(9) \quad \alpha = \frac{2}{\frac{1}{\alpha_1} + \frac{1}{\alpha_k}}$$

von α_1, α_k und betrachte statt $g(x)$ das Polynom

$$g^*(x) = \frac{g(x)}{(x+\alpha_1)(x+\alpha_k)} (x+\alpha)^2 = (x+\alpha)^2 (x+\alpha_2)(x+\alpha_3) \dots (x+\alpha_{k-1}).$$

Da die Werte $\alpha, \alpha, \alpha_2, \dots, \alpha_{k-1}$ nach (9) offenbar der Bedingung (8) genügen — also $g^*(x)$ zu der Schar der betrachteten Polynome gehört —, so wird der Satz bewiesen, wenn man zeigt, daß

$$(10) \quad \int_{-\beta}^2 g(x) (x-1) dx > \int_{-\beta}^2 g^*(x) (x-1) dx$$

gilt. Wir bilden die Differenz

$$(11) \quad \int_{-\beta}^2 g(x) (x-1) dx - \int_{-\beta}^2 g^*(x) (x-1) dx = \lambda \int_{-\beta}^2 \psi(x) (x-1) dx,$$

wobei

$$\lambda = \frac{(\alpha_1 - \alpha_k)^2}{\alpha_1 + \alpha_k},$$

$$\psi(x) = g(x) - g^*(x) = (x+\alpha_2)(x+\alpha_3) \dots (x+\alpha_{k-1})(x+\mu)$$

und

$$\mu = \frac{\alpha_1 \alpha_k}{\alpha_1 + \alpha_k}$$

ist. Es sei $\xi = \text{Min}(\beta, \mu)$, so wird

$$(12) \quad \int_{-\beta}^2 \psi(x)(x-1) dx \geq \int_{-\xi}^2 \psi(x)(x-1) dx.$$

Ist nämlich $\mu < \beta$, so ist $\int_{-\beta}^2 \psi(x)(x-1) dx > \int_{-\mu}^2 \psi(x)(x-1) dx$, da $\psi(x)(x-1)$ nach (8) im Intervall $(-\beta, -\mu)$ positiv ausfällt. Aus (11) und (12) folgt die Ungleichung:

$$(13) \quad \int_{-\beta}^2 g(x)(x-1) dx - \int_{-\beta}^2 g^*(x)(x-1) dx \geq \lambda \int_{-\xi}^2 \psi(x)(x-1) dx.$$

Die Wurzeln von $\psi(x)$ genügen wegen

$$\frac{1}{\alpha_2} + \frac{1}{\alpha_3} + \dots + \frac{1}{\alpha_{k-1}} + \frac{1}{\mu} = \frac{1}{\alpha_2} + \dots + \frac{1}{\alpha_{k-1}} + \frac{1}{\alpha_1} + \frac{1}{\alpha_k} = \gamma$$

der Bedingung (8). Da der Grad von $\psi(x)$ kleiner als k ist, ferner $\xi \leq \mu < \frac{k-1}{\gamma}$ und $\mu < \alpha_2$ ist, so gilt nach Voraussetzung der Satz für $\psi(x)$, d. h.

$$(14) \quad \int_{-\xi}^2 \psi(x)(x-1) dx > \int_{-\xi}^2 \left(x + \frac{k-1}{\gamma}\right)^{k-1} (x-1) dx.$$

Es ist aber

$$\begin{aligned} \int_{-\xi}^2 \left(x + \frac{k-1}{\gamma}\right)^{k-1} (x-1) dx &= \int_{x=-\xi}^{x=2} \frac{x-1}{k} \left(\frac{k-1}{\gamma} + x\right)^k - \frac{1}{k(k+1)} \left(\frac{k-1}{\gamma} + x\right)^{k+1} \\ &= \frac{k-1}{k(k+1)} \frac{\gamma-1}{\gamma} \left(\frac{k-1}{\gamma} + 2\right)^k + \frac{1}{k} \left(\frac{k-1}{\gamma} - \xi\right)^k \left\{1 + \xi \frac{k}{k+1} + \frac{k-1}{(k+1)\gamma}\right\} \end{aligned}$$

und aus dieser Formel folgt, da $\gamma \geq 1$ und $\xi < \frac{k-1}{\gamma}$ ist, die Ungleichung

$$\int_{-\xi}^2 \left(x + \frac{k-1}{\gamma}\right)^{k-1} (x-1) dx > 0.$$

Hieraus ergibt sich endlich nach (13) und (14) die Ungleichung

$$\int_{-\beta}^2 g(x)(x-1) dx - \int_{-\beta}^2 g^*(x)(x-1) dx > 0$$

und damit ist (10) bewiesen.

Hilfssatz II. Es seien $\alpha_1, \alpha_2, \dots, \alpha_k$ beliebige positive Zahlen und man bilde das Polynom

$$g(x) = (x + \alpha_1)(x + \alpha_2) \dots (x + \alpha_k).$$

Bedeutet γ eine feste positive Zahl, für welche $1 \leq \gamma$ ist, und betrachtet man alle Polynome, die den Bedingungen

$$\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_k} = \gamma, \quad \alpha_2 \geq \beta > 0$$

genügen, so erreicht das Integral

$$\int_{-\beta}^2 g(x) x^2(x-1) dx$$

sein Minimum für $\alpha_1 = \alpha_2 = \dots = \alpha_k = \frac{k}{\gamma}$, wenn $0 < \beta \leq \frac{k}{\gamma}$ ist.

Zum Beweis des Satzes wenden wir Induktion an, also setzen wir voraus, daß der Satz richtig ist, wenn der Grad von $g(x) < k$ ist. Sind die Werte α_i nicht alle gleich, so muß

$$\alpha_1 < \frac{k}{\gamma} < \alpha_k$$

sein. Wir führen nun auch hier, wie beim Beweis des Satzes I, das Polynom

$$g^*(x) = \frac{g(x)}{(x+\alpha_1)(x+\alpha_k)}, \quad \alpha = \frac{2}{\frac{1}{\alpha_1} + \frac{1}{\alpha_k}},$$

statt $g(x)$ ein und zeigen, daß die Ungleichung

$$\int_{-\beta}^2 g(x) x^2(x-1) dx > \int_{-\beta}^2 g^*(x) x^2(x-1) dx$$

gilt. Damit wird aber der Satz bewiesen. Man bilde die Differenz

$$(15) \quad \int_{-\beta}^2 g(x) x^2(x-1) dx - \int_{-\beta}^2 g^*(x) x^2(x-1) dx = \lambda \int_{-\beta}^2 \psi(x) x^2(x-1) dx,$$

wobei

$$\lambda = \frac{(\alpha_1 - \alpha_k)^2}{\alpha_1 + \alpha_k},$$

$$\psi(x) = (x + \alpha_2) \dots (x + \alpha_{k-1})(x + \mu), \quad \mu = \frac{\alpha_1 \alpha_k}{\alpha_1 + \alpha_k}$$

und

$$\frac{1}{\mu} + \frac{1}{\alpha_2} + \dots + \frac{1}{\alpha_{k-1}} = \gamma.$$

ist. Es sei $\xi = \min(\beta, \mu)$, so wird

$$\int_{-\beta}^2 \psi(x) x^2(x-1) dx \geq \int_{-\xi}^2 \psi(x) x^2(x-1) dx$$

und nach (15) folgt:

$$(16) \quad \int_{-\beta}^2 g(x) x^2(x-1) dx - \int_{-\beta}^2 g^*(x) x^2(x-1) dx \geq \lambda \int_{-\xi}^2 \psi(x) x^2(x-1) dx.$$

Da $\frac{1}{\mu} + \sum_{i=2}^{k-1} \frac{1}{\alpha_i} = \gamma$ und $\xi \leq \mu < \frac{k-1}{\gamma}$ ist, ferner $\mu < \alpha_2$, so gilt nach Voraussetzung der Satz für $\varphi(x)$, d. h.

$$(17) \quad \int_{-\xi}^2 \varphi(x) x^2(x-1) dx > \int_{-\xi}^2 \left(x + \frac{k-1}{\gamma}\right)^{k-1} x^2(x-1) dx.$$

Der Satz wird nun nach (15) bewiesen, wenn man noch zeigt, daß die rechte Seite von (17) positiv ist. Wir betrachten die Formel

$$(18) \quad \int_{-\xi}^2 \left(x + \frac{k-1}{\gamma}\right)^{k-1} x^2(x-1) dx = \int_{x=-\xi}^{x=2} \frac{x^3 - x^4}{k} \left(x + \frac{k-1}{\gamma}\right)^k + \\ + \frac{2x - 3x^2}{k(k+1)} \left(x + \frac{k-1}{\gamma}\right)^{k+1} + \frac{6x - 2}{k(k+1)(k+2)} \left(x + \frac{k-1}{\gamma}\right)^{k+2} - \\ - \frac{6}{k(k+1)(k+2)(k+3)} \left(x + \frac{k-1}{\gamma}\right)^{k+3} \\ = \int_{x=-\xi}^{x=2} F(x) = F(2) - F(-\xi).$$

Da

$$-F(-\xi) = \frac{\xi^3 + \xi^2}{k} \left(\frac{k-1}{\gamma} - \xi\right)^k + \frac{2\xi + 3\xi^2}{k(k+1)} \left(\frac{k-1}{\gamma} - \xi\right)^{k+1} + \\ + \frac{6\xi + 2}{k(k+1)(k+2)} \left(\frac{k-1}{\gamma} - \xi\right)^{k+2} + \frac{6}{k(k+1)(k+2)(k+3)} \left(\frac{k-1}{\gamma} - \xi\right)^{k+3} > 0$$

ist, muß man noch zeigen, daß $F(2) > 0$ gilt. Aus (18) ergibt sich für $F(2)$ die Formel

$$F(2) = \left(2 + \frac{k-1}{\gamma}\right)^k \frac{1}{k} \left\{4 - 8\left(\frac{2}{k+1} + \frac{k-1}{\gamma(k+1)}\right) + \right. \\ \left. + 10\frac{k+1}{k+2} \left(\frac{2}{k+1} + \frac{k-1}{\gamma(k+1)}\right)^2 - 6\frac{(k+1)^2}{(k+2)(k+3)} \left(\frac{2}{k+1} + \frac{k-1}{\gamma(k+1)}\right)^3\right\}.$$

Man bestätigt leicht, daß das Polynom

$$\varphi(x) = 4 - 8x + 10\frac{k+1}{k+2}x^2 - 6\frac{(k+1)^2}{(k+2)(k+3)}x^3$$

eine einzige positive Wurzel hat (es ist nämlich $\varphi'(x) < 0$, wenn $x > 0$ ist). Wegen $\gamma \geq 1$ folgt

$$\frac{2}{k+1} + \frac{k-1}{\gamma(k+1)} \leq 1$$

und hieraus, da $\varphi(1) > 0$ ist, ergibt sich

$$\varphi\left(\frac{2}{k+1} + \frac{k-1}{\gamma(k+1)}\right) > 0.$$

Infolgedessen ist $F(2)$ positiv, w. z. b. w.

§ 4.

T. Grünwald hat den folgenden Integralsatz bewiesen¹⁾: Es sei $f(x) = c \prod_{i=1}^n (x - a_i)$ ein Polynom mit lauter reellen Wurzeln: $a_1 > a_2 \geq a_3 \geq \dots \geq a_n$ und bedeute b_1 die im Intervall (a_2, a_1) liegende Wurzel von $f'(x)$. Ist $f(x)$ für $x > a_1$ positiv, so gilt die Ungleichung

$$\int_{a_2}^{a_1 + (a_1 - b_1)} f(x) dx \geq 0 \quad 2).$$

Dieser Satz folgt aus dem Hilfssatz I. Setzt man nämlich $(a_1 - b_1)x + b_1$ statt x , so gehen die Wurzeln $a_1 > a_2 \geq a_3 \geq \dots \geq a_n$ nacheinander in die Werte $1 > -\alpha_1 \geq -\alpha_2 \geq \dots \geq -\alpha_{n-1}$ über, wobei $\alpha_i > 0$ ist und die im Intervall $(-\alpha_1, 1)$ liegende Wurzel der ersten Ableitung gleich 0 wird. Man muß nun zeigen, daß das Integral

$$(19) \quad \int_{-\alpha_1}^1 g(x) (x-1) dx \geq 0$$

ist, wobei $g(x) = \prod_{i=1}^{n-1} (x + \alpha_i)$. Da $h'(0) = [(x-1)g(x)]' = 0$ ist, so folgt aus

$$\frac{h'(x)}{h(x)} = \frac{1}{x-1} + \sum_{i=1}^{n-1} \frac{1}{x+\alpha_i}$$

für $x = 0$ die Gleichung

$$(20) \quad \frac{1}{\alpha_1} + \frac{1}{\alpha_2} + \dots + \frac{1}{\alpha_{n-1}} = 1.$$

Also ist die Bedingung (8) für $\gamma = 1$, $k = n-1$ erfüllt. Da $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_{n-1}$ gilt, so folgt aus (20), daß $\alpha_1 \leq n-1$ ist, also nach dem Hilfssatz I ergibt sich die Ungleichung

$$\int_{-\alpha_1}^1 g(x) (x-1) dx \geq \int_{-\alpha_1}^1 (x+n-1)^{n-1} (x-1) dx.$$

Aus der Formel

$$\begin{aligned} \int_{-\alpha_1}^1 (x+n-1)^{n-1} (x-1) dx &= \left[\frac{(x+n-1)^n (x-2)}{n+1} \right]_{x=-\alpha_1}^{x=1} \\ &= \frac{1}{n+1} (n-1-\alpha_1)^n (\alpha_1+2) \geq 0 \end{aligned}$$

folgt endlich die Richtigkeit von (19).

¹⁾ T. Grünwald, Über Polynome mit reellen Nullstellen (ungarisch). Mat. fizik. Lapok 46 (1939), S. 31–57.

²⁾ Aus diesem Satz leitete T. Grünwald den folgenden Satz von I. Schur ab: Wenn ein Polynom nur reelle Wurzeln hat, so ist der Abstand zwischen der größten Wurzel des Polynoms und der größten Wurzel der ersten Ableitung nicht größer als der Abstand der größten Wurzeln der ersten und zweiten Ableitungen.

Sind die Wurzeln von $f(x)$: $a_1 \geq a_2 > a_3 > a_4 \geq \dots \geq a_n$ und ist $f(x)$ für $x > a_1$ positiv, d. h. $f(x) = \prod_{i=1}^n (x + \alpha_i)$, so ist $f(x)$ im Intervall (a_3, a_2) positiv und wir können den Satz von T. Grünwald folgenderweise erweitern: Setzen wir $(a_3 - b_3)x + b_3$ statt x , wobei b_3 die Wurzel von $f'(x)$ in (a_4, a_3) bedeutet, dann gehen die Wurzeln von $f(x)$: $a_1 \geq a_2 > a_3 > a_4 \geq \dots \geq a_n$ in die Werte: $x_2 \geq x_1 > 1 > -\alpha_1 \geq -\alpha_2 \geq \dots \geq -\alpha_{n-3}$, $\alpha_i > 0$, nacheinander über. Die im Intervall $(-\alpha_1, 1)$ liegende Wurzel der Ableitung wird gleich Null. Setzen wir

$$\begin{aligned} \varphi(x) &= (x - x_1)(x - x_2)(x - 1)(x + \alpha_1)(x + \alpha_2) \dots (x + \alpha_{n-3}) \\ &= (x - x_1)(x - x_2)(x - 1)g(x), \end{aligned}$$

so gilt der

Satz 3. Genügen die Wurzeln x_1 und x_2 der Bedingung

$$(21) \quad x_2 \geq x_1 \geq 2 + 2 \frac{\int_{-a_1}^2 g(x)(x-1)^2 dx}{\int_{-a_1}^2 g(x)(x-1) dx},$$

so ist

$$\int_{-a_1}^2 \varphi(x) dx > 0.$$

Beweis. Wir betrachten das Integral

$$\begin{aligned} (22) \quad \int_{-a_1}^2 \varphi(x) dx &= \int_{-a_1}^2 g(x)(x-1)(x-x_1)(x-x_2) dx \\ &= \int_{-a_1}^2 g(x)x^2(x-1) dx - (x_1+x_2) \int_{-a_1}^2 g(x)x(x-1) dx + \\ &\quad + x_1x_2 \int_{-a_1}^2 g(x)(x-1) dx. \end{aligned}$$

Wir zeigen erstens, daß das Integral $\int_{-a_1}^2 g(x)x^2(x-1) dx$ — in der rechten Seite — positiv ist. Da $\varphi'(0) = 0$ gilt, so folgt für $x = 0$ aus

$$\frac{\varphi'(x)}{\varphi(x)} = \frac{1}{x-x_1} + \frac{1}{x-x_2} + \frac{1}{x-1} + \sum_{i=1}^{n-3} \frac{1}{x+\alpha_i},$$

daß die Wurzeln von $\varphi(x)$ der Bedingung

$$\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_{n-3}} = 1 + \frac{1}{x_1} + \frac{1}{x_2}$$

genügen. Setzt man $\gamma = 1 + \frac{1}{x_1} + \frac{1}{x_2}$, so genügen die α_i der Bedingung

$$(23) \quad \frac{1}{\alpha_1} + \frac{1}{\alpha_2} + \dots + \frac{1}{\alpha_{n-3}} = \gamma, \quad \gamma \geq 1.$$

Da $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_{n-3}$ sind, genügt α_1 der Bedingung $\alpha_1 \leq \frac{n-3}{\gamma}$ und darum sind alle Bedingungen des Hilfssatzes II erfüllt. Das Integral

$$\int_{-\alpha_1}^2 g(x) x^2(x-1) dx \text{ erreicht also sein Minimum für } \alpha_1 = \alpha_2 = \dots = \alpha_{n-3} \\ = \frac{n-3}{\gamma}, \text{ d. h.}$$

$$\int_{-\alpha_1}^2 g(x) x^2(x-1) dx \geq \int_{-\alpha_1}^2 \left(x + \frac{n-3}{\gamma}\right)^{n-3} x^2(x-1) dx.$$

Die rechte Seite dieser Ungleichung ist aber positiv, wie wir es beim Beweis des Hilfssatzes II gezeigt haben [man setze in (18) $k = n-2$].

Jetzt muß man noch zeigen, daß in (22) die Differenz

$$-(x_1 + x_2) \int_{-\alpha_1}^2 g(x) x(x-1) dx + x_1 x_2 \int_{-\alpha_1}^2 g(x) (x-1) dx$$

positiv ausfällt, wenn x_1 der Bedingung (21) genügt. Da $x_1 \leq x_2$ gilt, ist diese Differenz positiv, wenn

$$(24) \quad x_1 > 2 \frac{\int_{-\alpha_1}^2 g(x) x(x-1) dx}{\int_{-\alpha_1}^2 g(x) (x-1) dx}$$

gilt. Aus der Formel

$$\int_{-\alpha_1}^2 g(x) x(x-1) dx - \int_{-\alpha_1}^2 g(x) (x-1) dx = \int_{-\alpha_1}^2 g(x) (x-1)^2 dx$$

folgt aber:

$$\frac{\int_{-\alpha_1}^2 g(x) x(x-1) dx}{\int_{-\alpha_1}^2 g(x) (x-1) dx} = 1 + \frac{\int_{-\alpha_1}^2 g(x) (x-1)^2 dx}{\int_{-\alpha_1}^2 g(x) (x-1) dx}$$

und hiernach ist (24) erfüllt, wenn x_1 der Bedingung (21) genügt.

Für den in (21) vorkommenden Bruch bestimmen wir endlich eine obere Schranke. Nach dem Satz der arithmetischen und geometrischen Mittel ist:

$$g(x) = \alpha_1 \alpha_2 \dots \alpha_{n-3} \left(\frac{x}{\alpha_1} + 1\right) \left(\frac{x}{\alpha_2} + 1\right) \dots \left(\frac{x}{\alpha_{n-3}} + 1\right) \\ \leq \alpha_1 \alpha_2 \dots \alpha_{n-3} \left\{ \frac{\left(\frac{x}{\alpha_1} + 1\right) + \left(\frac{x}{\alpha_2} + 1\right) + \dots + \left(\frac{x}{\alpha_{n-3}} + 1\right)}{n-3} \right\}^{n-3}$$

und hieraus folgt nach (23):

$$g(x) \leq \alpha_1 \alpha_2 \dots \alpha_{n-3} \left(1 + x \frac{\gamma}{n-3}\right)^{n-3}$$

Infolgedessen gilt die Ungleichung

$$\begin{aligned} (25) \quad \int_{-\alpha_1}^2 g(x) (x-1)^2 dx &\leq \prod_{i=1}^{n-3} \alpha_i \left(\frac{\gamma}{n-3}\right)^{n-3} \int_{-\alpha_1}^2 \left(\frac{n-3}{\gamma} + x\right)^{n-3} (x-1)^2 dx \\ &\leq \prod_{i=1}^{n-3} \alpha_i \left(\frac{\gamma}{n-3}\right)^{n-3} \int_{\frac{n-3}{\gamma}}^2 \left(\frac{n-3}{\gamma} + x\right)^{n-3} (x-1)^2 dx, \end{aligned}$$

weil $\alpha_1 \leq \frac{n-3}{\gamma}$ und $\left(\frac{n-3}{\gamma} + x\right)^{n-3} (x-1)^2$ für $x \geq -\frac{n-3}{\gamma}$ immer positiv ist. Aus der Formel

$$\begin{aligned} &\int \left(\frac{n-3}{\gamma} + x\right)^{n-3} (x-1)^2 dx \\ &= (x-1)^2 \frac{\left(\frac{n-3}{\gamma} + x\right)^{n-2}}{n-2} - \frac{2(x-1)}{(n-2)(n-1)} \left(\frac{n-3}{\gamma} + x\right)^{n-1} + 2 \frac{\left(\frac{n-3}{\gamma} + x\right)^n}{(n-2)(n-1)n} \end{aligned}$$

ergibt sich nun leicht, da $\gamma \geq 1$ ist, die Ungleichung

$$(26) \quad \int_{-\frac{n-3}{\gamma}}^2 \left(\frac{n-3}{\gamma} + x\right)^{n-3} (x-1)^2 dx < \frac{(n-1)^{n-2}}{n}.$$

Nach (23), (25) und (26) gewinnt man also die Abschätzung

$$(27) \quad \int_{-\alpha_1}^2 g(x) (x-1)^2 dx \leq \left(\frac{\sum_{i=1}^{n-3} \frac{1}{\alpha_i}}{n-3}\right)^{n-3} \frac{(n-1)^{n-2}}{n} \prod_{i=1}^{n-3} \alpha_i.$$

Für den Nenner des in (21) vorkommenden Bruches folgt nach dem Hilfsatz I die untere Schranke

$$(28) \quad \int_{-\alpha_1}^2 g(x) (x-1) dx \geq \int_{-\alpha_1}^2 \left(x + \frac{n-3}{\gamma}\right)^{n-3} (x-1) dx.$$

Aus der Formel

$$\int \left(x + \frac{n-3}{\gamma}\right)^{n-3} (x-1) dx = \left(\frac{n-3}{\gamma} + x\right)^{n-2} \frac{1}{n-2} \left(x \frac{n-2}{n-1} - 1 - \frac{n-3}{\gamma(n-1)}\right)$$

ergibt sich leicht, daß die rechte Seite von (28) die untere Schranke

$$\int_{-\alpha_1}^2 \left(x + \frac{n-3}{\gamma}\right)^{n-3} (x-1) dx > \frac{1}{n-2} \left(\frac{n-3}{\gamma} - \alpha_1\right)^{n-2}$$

hat. Also gilt nach (23) und (28) die Ungleichung

$$(29) \quad \int_{-\alpha_1}^2 g(x)(x-1) dx > \frac{1}{n-2} \left(\frac{n-4 - \sum_{i=2}^{n-3} \frac{\alpha_i}{\alpha_i}}{\sum_{i=1}^{n-3} \frac{1}{\alpha_i}} \right)^{n-2}$$

Nach (27) und (29) gewinnt man endlich für den in (21) vorkommenden Bruch die Abschätzung:

$$\frac{\int_{-\alpha_1}^2 g(x)(x-1)^2 dx}{\int_{-\alpha_1}^2 g(x)(x-1) dx} < \frac{(n-2)(n-1)^{n-2}}{n(n-1)^{n-3}} \left(\sum_{i=1}^{n-3} \frac{1}{\alpha_i} \right)^{2n-5} \prod_{i=1}^{n-3} \alpha_i \left(n-4 - \sum_{i=2}^{n-3} \frac{\alpha_i}{\alpha_i} \right)^{2-n}.$$

(Eingegangen am 1. 10. 1941.)

Eine Methode zur Verallgemeinerung der gewöhnlichen Differentialgleichungen.

Von

Josefa von Schwarz in Berlin.

Einleitung.

Wir zeigen im folgenden die Lösbarkeit einer Klasse von Problemen, die zwanglos als eine Verallgemeinerung der gewöhnlichen Differentialgleichungen gelten können.

Man kommt zu dieser Verallgemeinerung, indem man zunächst die Frage aufwirft, inwieweit bei der Formulierung des Problems der gewöhnlichen Differentialgleichungen bzw. bei dem Nachweis der Existenz von Lösungen desselben von der Voraussetzung Gebrauch gemacht wird, daß dieses Problem — wie üblich — in einem euklidischen Raum gestellt wird. Es ist leicht einzusehen, daß dies nur in geringem Maße der Fall ist; es wird lediglich die Eigenschaft des euklidischen Raumes benutzt, ein halbmétrischer Raum mit gleichmäßig stetigen Abständen zu sein.

Zu dem Zweck formulieren wir das Problem der gewöhnlichen Differentialgleichungen folgendermaßen:

In jedem Punkt r eines Gebietes eines zunächst als euklidisch vorausgesetzten Raumes beginne ein glatter — d. h. mit stetigen Tangenten versehener Bogen C_r . Verlangt wird, durch jeden Punkt des Gebietes eine Kurve zu legen, die in jedem ihrer Punkte r den zugehörigen Bogen C_r berührt. Das bekannte Existenztheorem besagt, daß dies immer möglich ist, wenn in einer genügend kleinen Umgebung jedes Punktes r des Gebietes irgend zwei dort beginnende Bögen C_r und $C_{r'}$ beliebig wenig voneinander abweichende Tangenten haben.

Diese Bedingung läßt sich offenbar auch so schreiben:

Sind p' und q' bzw. p'' und q'' zwei in der genügend klein gewählten Umgebung von r aufeinanderfolgende Punkte von C_r bzw. $C_{r'}$, so ist $\cos(p', q'; p'', q'')$, d. h. der Cosinus des Winkels zwischen den durch p', q' bzw. p'', q'' bestimmten Strecken, beliebig nahe an 1.

Beachten wir aber nun, daß

$$\cos(p', q'; p'', q'') = \frac{-\delta(p', p'')^2 - \delta(q', q'')^2 + \delta(p', q'')^2 + \delta(q', p'')^2}{2\delta(p', q') \cdot \delta(p'', q'')}$$

sich nur durch die — zunächst als euklidisch vorausgesetzten — Abstände $\delta(p', p'')$ usw. zwischen den vier Punkten ausdrücken läßt, so liegt eine Verallgemeinerung dieser Funktion $\cos(p', q'; p'', q'')$ auf Punktepaare in halbmetrischen Räumen auf der Hand. Man braucht sich in der obigen Formel nur die euklidischen Abstände durch die Abstände in dem betreffenden Raum ersetzt zu denken.

Mit Hilfe der so definierten Funktion $\cos(p', q'; p'', q'')$ läßt sich in einem halbmetrischen Raum mit gleichmäßig stetigen Abständen analog wie im euklidischen Raum definieren, wann zwei Kurven sich berühren. Auf diese Weise läßt sich das Problem der gewöhnlichen Differentialgleichungen, wie wir es oben formulierten, wörtlich in einen halbmetrischen Raum mit gleichmäßig stetigen Abständen übertragen, wobei es sich herausstellt, daß die Eigenschaften der Funktion $\cos(p', q'; p'', q'')$ in einem solchen Raum auch ausreichen, das erwähnte Existenztheorem zu beweisen.

Dadurch wird aber eine weitere Verallgemeinerung des Differentialgleichungsproblems nahegelegt. Wir führen nämlich Funktionen $\alpha(p', q'; p'', q'')$ ein, die gerade durch die Eigenschaften von $\cos(p', q'; p'', q'')$ charakterisiert sind, die wir benötigen, um das Existenztheorem zu beweisen. Die Funktion \cos ist also ein Spezialfall der α -Funktionen. Daß die Auswahlmöglichkeiten für die α -Funktionen auch im euklidischen Raum sehr groß sind, beweist das in den §§ 16ff. durchgeführte Beispiel.

Definiert man nun analog wie mit Hilfe der Funktion $\cos(p', q'; p'', q'')$, wann sich zwei Kurven α -berühren, so kommt man zu der eingangs erwähnten Klasse von Problemen, die wir kurz als α -Probleme bezeichnen, und für die das Existenztheorem dann offenbar auch gültig ist.

Bei der ausführlichen Ableitung dieser Ergebnisse auf folgenden Seiten haben wir, um Wiederholungen zu vermeiden, nicht den eben geschilderten Gedankengang entwickelt, sondern — nach Voranstellung einiger Sätze über Polygone und Kurven in einem halbmetrischen Raum mit gleichmäßig stetigen Abständen — sogleich die Funktionen $\alpha(p', q'; p'', q'')$ axiomatisch eingeführt und das Existenztheorem allgemein bewiesen.

Nachträglich wird dann gezeigt, daß $\cos(p', q'; p'', q'')$ eine α -Funktion ist. Schließlich beleuchten wir die Brauchbarkeit dieser Begriffsbildung auch im euklidischen Raum, indem wir durch Wahl einer passenden α -Funktion einen neuen Beweis für die Existenz von Lösungen der Paratingent-Gleichungen¹⁾ des Herrn Zaremba erbringen.

¹⁾ St. Ch. Zaremba, Sur les équations au paratingent. Bull. d. Sc. Math. 60 (1936), S. 139—160.

I. Allgemeines über Polygone und Streckenbilder in halbmétrischen Räumen mit gleichmäßig stetigen Abständen²⁾.

§ 1. Wir legen unseren Untersuchungen einen Punktraum R zugrunde, in dem zu je zwei Punkten p, q ein Abstand $\delta(p, q)$ definiert ist, der folgenden Bedingungen genügt:

$$1) \delta(p, q) = \delta(q, p) > 0 \quad \text{für } q \neq p$$

$$\delta(p, p) = 0,$$

$$2) \text{ zu jedem } \varepsilon > 0 \text{ existiert ein } \Delta > 0, \text{ so daß aus}$$

$$\delta(p, p') + \delta(q, q') < \Delta$$

stets

$$|\delta(p, q') - \delta(p', q)| < \varepsilon$$

folgt.

Einen solchen Raum nennt man einen *halbmétrischen Raum* (Bedingung 1) mit *gleichmäßig stetigen Abständen* (Bedingung 2).

Polygon nennen wir eine geordnete endliche Menge von Punkten von R und schreiben: $P = \{p_1, \dots, p_m\}$. Seine *Länge* ist

$$\lambda(P) = \sum_{i=1}^{m-1} \delta(p_i, p_{i+1}).$$

Außerdem werden die Abkürzungen

$$\text{Spanne: } \mu(P) = \delta(p_1, p_m),$$

$$\text{Durchmesser: } D(P) = \max_{\substack{i=1, \dots, m-1 \\ j=2, \dots, m}} (\delta(p_i, p_j)),$$

$$\text{Norm: } \nu(P) = \max_{i=1, \dots, m-1} (\delta(p_i, p_{i+1}))$$

verwendet³⁾.

Als (stetiges)⁴⁾ *Streckenbild* $C = p[\alpha, \beta]$ bezeichnen wir eine stetige Zuordnung eines Punktes $p(\gamma)$ zu jeder Zahl γ des abgeschlossenen Intervalls $[\alpha, \beta]$, ($\beta > \alpha$). Als *Teilpolygon* von $C = p[\alpha, \beta]$ werde jedes Polygon $P = \{p(\gamma_1), p(\gamma_2), \dots, p(\gamma_m)\}$ bezeichnet, für welches $\alpha = \gamma_1 < \gamma_2 < \dots < \gamma_{m-1} < \gamma_m = \beta$ ist. Wir nennen *C-Norm* von P und bezeichnen mit $\nu_C(P)$ die größte der Zahlen $\gamma_{i+1} - \gamma_i$. Wir bezeichnen als *obere* bzw. *untere Länge* von C die obere bzw. untere Schranke der Zahlen λ , für welche eine Folge

²⁾ Zu der im folgenden angewandten Bezeichnungswiese vgl. man K. Menger, Die métrische Methode in der Variationsrechnung. Ergebn. eines math. Kolloquiums, Heft 8, §§ 1–5.

³⁾ Es bedeutet keine Beschränkung der Allgemeinheit, nur solche Polygone zu betrachten, für die stets $\delta(p_i, p_{i+1}) \neq 0$, ($i = 1, \dots, m-1$) ist, was wir im folgenden tun wollen.

⁴⁾ Stetige Streckenbilder nennen wir auch kurz *Kurven*.

P_1, P_2, \dots von Teilpolygonen von C mit $\lim_{n \rightarrow \infty} v_C(P_n) = 0$ und $\lim_{n \rightarrow \infty} \lambda(P_n) = \lambda$ existiert. Stimmen obere und untere Länge von C überein, so nennen wir ihren gemeinsamen (endlichen oder unendlichen) Wert die *Länge* von C .

§ 2. Wir bezeichnen ein stetiges Streckenbild $C = p[\alpha, \beta]$ als *Häufungskurve* einer Gesamtheit von Polygonen P' , wenn zu jedem $\Delta > 0$ und zu jedem Teilpolygon $P_C = \{p(\gamma_1), \dots, p(\gamma_m)\}$ von C mindestens ein Polygon P' der Gesamtheit existiert, für welches $\delta(P', P_C) < \Delta^5$ gilt, d. h. ein Polygon $P' = \{p'_1, \dots, p'_m\}$ mit folgenden Eigenschaften:

P' zerfällt in $m-1$ Abschnitte $A'^i = \{p'_{h_i}, \dots, p'_{h_{i+1}}\}$, ($i = 1, \dots, m-1$) mit $h_1 = 1$ und $h_m = m'$, und es gilt $\delta(p'_{h_i}, p_i) < \Delta$, $\delta(p'_m, p_m) < \Delta$ und $A'^i \subset U(C_i, \Delta)^7$, wobei $C_i = p[\gamma_i, \gamma_{i+1}]$ ist.

Wir sagen, daß eine Gesamtheit von Polygonen gegen eine Kurve C konvergiert, wenn es zu jedem $v^* > 0$ ein P' gibt, so daß $v(P') < v^*$ ist, und wenn man zu jedem $\Delta > 0$ und $v_C^* > 0$ ein $v^* > 0$ angeben kann, so daß es zu jedem P' , für das $v(P') < v^*$ gilt, mindestens ein Teilpolygon P_C von C mit $v_C(P_C) < v_C^*$ gibt, derart, daß $\delta(P', P_C) < \Delta$.

Das Streckenbild, gegen das eine Gesamtheit von Polygonen konvergiert, ist offenbar eine Häufungskurve (und auch die einzige) desselben. Andererseits enthält eine Gesamtheit von Polygonen, die ein Streckenbild als Häufungskurve besitzt, stets eine Teilgesamtheit, die gegen dieses Streckenbild konvergiert.

§ 3. Es gelten folgende Sätze, deren einfachen Beweis wir übergehen.

Satz 1. In einem halbmétrischen Raum mit gleichmäßig stetigen Abständen sei eine Folge von Polygonen P'_1, P'_2, \dots mit folgenden Eigenschaften gegeben:

1) $\lambda(P'_n) \leq \bar{\lambda}^*$ ($n = 1, 2, \dots$).

2) Es existiere eine Zahl $\Theta_1 \geq 0$ und eine Zahl $D_1 > 0$, so daß für jeden Abschnitt A'_n von P'_n , für den $D(A'_n) < D_1$ ist, stets

$$(1) \quad \mu(A'_n) \leq \sqrt{1 + \Theta_1} \cdot \lambda(A'_n)$$

ist.

Dann gilt für die obere Länge $\bar{\lambda}(C)$ jeder Häufungskurve C dieser Polygonfolge.

$$\bar{\lambda}(C) \leq \sqrt{1 + \Theta_1} \cdot \bar{\lambda}^*.$$

⁵⁾ Zu dieser Begriffsbildung vgl. K. Menger, l. c. S. 10.

⁶⁾ Unter einem Abschnitt eines Polygons verstehen wir ein Polygon, das aus lauter unmittelbar aufeinanderfolgenden Punkten jenes Polygons besteht.

⁷⁾ Ist M irgendeine Punktmenge von R , so bezeichnen wir mit $U(M, \Delta)$ die Gesamtheit aller Punkte von R , die von mindestens einem Punkt von M eine Entfernung $< \Delta$ haben. $M \subset M'$ bedeutet wie üblich, daß die Menge M in der Menge M' enthalten ist.

Satz 2. In einem halbmétrischen Raum R mit gleichmäßig stetigen Abständen sei eine Folge von Polygonen P_1, P_2, \dots mit folgenden Eigenschaften gegeben:

$$1) \lambda(P_n) \geq \lambda^* \quad (n = 1, 2, \dots).$$

2) Es existiere eine Zahl Θ_2 , ($0 \leq \Theta_2 < 1$) und eine Zahl $D_2 > 0$, so daß für jeden Abschnitt A'_n von P_n , für den $D(A'_n) < D_2$ ist, stets

$$\sqrt{1 - \Theta_2} \cdot \lambda(A'_n) \leq \mu(A'_n)$$

ist.

Dann gilt für die untere Länge $\lambda(C)$ jeder Häufungskurve C dieser Polygonfolge

$$\sqrt{1 - \Theta_2} \cdot \lambda^* \leq \lambda(C).$$

Satz 3. In einem halbmétrischen Raum R mit gleichmäßig stetigen Abständen sei eine Folge von Polygonen P_1, P_2, \dots mit folgenden Eigenschaften gegeben: Es existieren Zahlen Θ_1 und Θ_2 , ($0 \leq \Theta_1$, $0 \leq \Theta_2 < 1$), so daß für jeden Abschnitt A'_n von P_n stets

$$\sqrt{1 - \Theta_2} \cdot \lambda(A'_n) \leq \mu(A'_n) \leq \sqrt{1 + \Theta_1} \cdot \lambda(A'_n)$$

ist.

Dann ist jede Häufungskurve C dieser Polygonfolge ein Bogen.

§ 4. Satz. Ist der Raum R kompakt und $\lim_{n \rightarrow \infty} v(P_n) = 0$, so folgt unter den Voraussetzungen von Satz 1 von § 3, daß immer wenigstens eine Häufungskurve der Polygonfolge existiert.

Beweis. Zum Zweck einer übersichtlichen Bezeichnungsweise führen wir zunächst eine Abbildung des Intervalls $[0 \leq \zeta \leq 1]$ auf irgendein Polygon $P = \{p_1, \dots, p_m\}$ ein. Wir ordnen nämlich jedem Wert ζ dieses Intervalls den eindeutig bestimmten Punkt $p(\zeta) = p_j$ von P zu, für den

$$\zeta - \frac{1}{\lambda(P)} \cdot \sum_{i=1}^{j-1} \delta(p_i, p_{i+1}) \quad (j = 1, \dots, m)$$

den kleinsten nicht negativen Wert hat⁸⁾.

Bezeichnen wir mit ζ^* und ζ^{**} zwei Werte, für die $0 \leq \zeta^* \leq \zeta^{**} \leq 1$ gilt, mit $P(\zeta^*, \zeta^{**})$ den Abschnitt von P , der mit $p(\zeta^*)$ beginnt und mit $p(\zeta^{**})$ endet, so gilt offenbar

$$(2) \quad \lambda(P) \cdot (\zeta^{**} - \zeta^*) + v(P) > \lambda(P(\zeta^*, \zeta^{**})).$$

⁸⁾ $\sum_{i=1}^0 \delta(p_i, p_{i+1})$ wird gleich Null gesetzt.

Ferner beachten wir, daß sich aus (1), da $\lim_{n \rightarrow \infty} \nu(P'_n) = 0$ ist, ohne weiteres folgendes ergibt:

Zu jedem $\varepsilon > 0$ kann man n so groß wählen, daß für $D(A'_n) \geq D_1$ stets

$$(3) \quad \sqrt{1 + \Theta_1} \cdot \lambda(A'_n) > D_1 - \varepsilon$$

ist.

Aus (2) und (3) folgt aber:

Man kann ein $N > 0$ und ein $\Delta > 0$ angeben, so daß für $n > N$ und $\zeta^{**} - \zeta^* < \Delta$ stets

$$(4) \quad \bar{\lambda} \cdot \Delta + \nu(P'_n) > \frac{\delta(p'_n(\zeta^*), p'_n(\zeta^{**}))}{\sqrt{1 + \Theta_1}}$$

gilt.

Nun wählen wir nach dem bekannten Hilbertschen Verfahren aus der Polygonfolge P_1, P'_2, \dots eine sogenannte *Diagonalfolge* $P_{11}, P_{22}, \dots, P_{nn}, \dots$ von der Art, daß für jede diadisch rationale Zahl ζ , ($0 \leq \zeta \leq 1$) die Punkte $p'_{11}(\zeta), p'_{22}(\zeta), \dots, p'_{nn}(\zeta), \dots$ gegen einen Punkt $p(\zeta)$ des Raumes R konvergieren. Diese Zuordnung ist, wie wir zeigen wollen, gleichmäßig stetig, d. h. man kann zu jedem $\varepsilon > 0$ ein $\Delta > 0$ angeben, so daß, wenn für die diadisch rationalen Zahlen ζ^* und ζ^{**} , ($0 \leq \zeta^* \leq \zeta^{**} \leq 1$) die Ungleichung $\zeta^{**} - \zeta^* < \Delta$ gilt, dann auch immer $\delta(p(\zeta^*), p(\zeta^{**})) < \varepsilon$ ist.

Man kann zunächst wegen der gleichmäßigen Stetigkeit der Abstände zu $\varepsilon > 0$ ein $\Delta_1 > 0$ finden, so daß, wenn

$$(5) \quad \delta(p'_{nn}(\zeta^*), p(\zeta^*)) + \delta(p'_{nn}(\zeta^{**}), p(\zeta^{**})) < \Delta_1, \\ |\delta(p'_{nn}(\zeta^*), p'_{nn}(\zeta^{**})) - \delta(p(\zeta^*), p(\zeta^{**}))| < \frac{\varepsilon}{2}$$

ist.

Es kommt also nur darauf an, n so zu bestimmen, daß sowohl (5) gilt, als auch

$$(6) \quad \delta(p'_{nn}(\zeta^*), p'_{nn}(\zeta^{**})) < \frac{\varepsilon}{2}.$$

(5) ist nun gemäß der Definition der Diagonalfolge $P_{11}, P_{22}, \dots, P_{nn}, \dots$ durch Wahl eines genügend großen n erfüllbar. Wegen der Ungleichung (4) können wir n so groß und damit $\nu(P'_{nn})$ so klein, sowie auch Δ so klein wählen, daß auch (6) erfüllt ist.

Die Zuordnung der diadisch rationalen Zahlen ζ zu den Punkten $p(\zeta)$ ist also gleichmäßig stetig. Es ist daher bekanntlich möglich, diese Zuordnung zu einer stetigen Abbildung des Intervalls ($0 \leq \zeta \leq 1$), d. h. zu einem stetigen Streckenbild $C = p[0, 1]$ zu erweitern.

Dieses Streckenbild ist eine Häufungskurve der Folge $P_1, P'_2, \dots, P'_n, \dots$, genauer: die Diagonalfolge $P_{11}, P_{22}, \dots, P_{nn}, \dots$ konvergiert gegen C .

Dazu brauchen wir nach § 2 nur zu zeigen, daß es zu jedem $\Delta > 0$ und $\nu_C^* > 0$ ein $N > 0$ gibt, so daß es zu jedem P'_{nn} ($n > N$) mindestens ein Teilpolygon P_C von C mit $\nu_C(P) < \nu_C^*$ gibt, so daß $\delta(P'_{nn}, P_C) < \Delta$ ist.

Wegen der gleichmäßigen Stetigkeit der Abstände gibt es — wie immer die drei Punkte $p'_{nn}(\zeta_j)$, $p(\zeta_j)$, $p'_{nn}(\zeta)$ gewählt werden — zu jedem $\Delta > 0$ ein $\Delta' > 0$, so daß aus

$$\delta(p'_{nn}(\zeta_j), p'_{nn}(\zeta)) < \Delta', \quad \delta(p'_{nn}(\zeta_j), p(\zeta_j)) < \frac{\Delta}{2}$$

stets folgt

$$\delta(p'_{nn}(\zeta), p(\zeta_j)) < \Delta.$$

Man kann nun nach (4) zu jedem $\Delta' > 0$ die Zahl $N' > 0$ so groß und $\Delta'' > 0$ so klein wählen, daß $\delta(p'_{nn}(\zeta_j), p'_{nn}(\zeta)) < \Delta'$ gilt für irgend zwei Punkte $p'_{nn}(\zeta_j)$, $p'_{nn}(\zeta)$ mit $|\zeta - \zeta_j| < \Delta''$ eines Polygons P'_{nn} ($n > N'$) der Diagonalfolge.

Wir wählen nun eine *endliche* Folge diadisch rationaler Zahlen $\zeta_1, \zeta_2, \dots, \zeta_k$, für die $\zeta_1 = 0$, $\zeta_k = 1$, $\zeta_{i+1} - \zeta_i < \min(\nu_C^*, \Delta'')$, ($i = 1, \dots, k-1$) ist.

Wir können schließlich nach Definition der Diagonalfolge eine Zahl $N'' > 0$ finden, so daß

$$\delta(p'_{nn}(\zeta_j), p(\zeta_j)) < \frac{\Delta}{2} \quad (j = 1, \dots, k, n > N'')$$

ist.

Wählen wir also $N \geq \max(N', N'')$, so ist $P_C = \{p(\zeta_1), \dots, p(\zeta_k)\}$ ein Teilpolygon P von C , für das $\nu_C(P) < \nu_C^*$ und $\delta(P'_{nn}, P_C) < \Delta$ (für $n > N$) gilt.

I. Einführung der Funktionen $\kappa(p, q; p', q')$ und Beweis des Existenztheorems für die κ -Probleme.

§ 5. Zu je zwei Punktepaaren p, q und p', q' ($p \neq q, p' \neq q'$) eines halbmétrischen Raumes R sei eine reelle Zahl $\kappa(p, q; p', q')$ zugeordnet, die den folgenden Bedingungen unterworfen ist.

$$\text{I. } \kappa(p, q; p', q') = \kappa(p', q'; p, q).$$

II. Zu je zwei Punktepaaren p, q und p', q' und jedem $\varepsilon > 0$ gibt es ein $\Delta > 0$, so daß aus

$$\delta(\bar{p}, \bar{p}) + \delta(q, \bar{q}) < \Delta$$

folgt

$$|\kappa(p, q; p', q') - \kappa(\bar{p}, \bar{q}; p', q')| < \varepsilon.$$

III. Es gibt zu jedem $\Theta > 0$ und jedem Punkt p von R ein $\vartheta > 0$ und $\varrho > 0$, so daß wenn immer für zwei beliebige in der Umgebung $U(p, \varrho)$ liegenden Polygone $P = \{p_1, \dots, p_m\}$ und $P' = \{p'_1, \dots, p'_m\}$,

$$1 - \vartheta \leq \left\{ \begin{array}{l} \kappa(p_i, p_{i+1}; p_j, p_{j+1}) \\ \kappa(p'_i, p'_{i+1}; p'_j, p'_{j+1}) \end{array} \right\} \leq 1 + \vartheta \quad \left(\begin{array}{l} i, j = 1, \dots, m-1 \\ k, l = 1, \dots, m'-1 \end{array} \right)$$

gilt, stets

$$1 - \Theta \leq \kappa(p_1, p_m; p'_1, p'_m) \leq 1 + \Theta^0$$

ist.

IV. Es gibt eine Zahl $\vartheta^* > 0$ und eine Zahl Θ^* , ($0 < \Theta^* < 1$), so daß für jedes Polygon $P = \{p_1, \dots, p_m\}$, für das

$$1 - \vartheta^* \leq \kappa(p_i, p_{i+1}; p_j, p_{j+1}) \leq 1 + \vartheta^*$$

ist, stets

$$\lambda(P) \cdot \sqrt{1 - \Theta^*} \leq \mu(P) \leq \sqrt{1 + \Theta^*} \cdot \lambda(P)$$

gilt.

§ 6. Ist eine bestimmte Funktion $\kappa(p, q; p', q')$ gemäß den Bedingungen des § 5 definiert, so sagen wir, daß zwei stetige Streckenbilder $C = p[\alpha, \beta]$ und $C' = p'[\alpha', \beta']$ sich im Punkte p_0 , der für beide Streckenbilder ein einfacher¹⁰⁾ Punkt sei — im Sinne dieser κ -Funktion —, κ -berühren, falls es zu jedem $\varepsilon > 0$ ein $\varrho > 0$ gibt, so daß aus

$$p, q \subset C \cdot U(p_0, \varrho); \quad p', q' \subset C' \cdot U(p_0, \varrho) \quad (p \xrightarrow{C} q, \quad p' \xrightarrow{C'} q')^{11)} \\ (p \neq q, \quad p' \neq q')$$

stets

$$1 - \varepsilon \leq \kappa(p, q; p', q') \leq 1 + \varepsilon$$

folgt¹²⁾.

Ein Streckenbild, das sich in einem seiner (einfachen) Punkte selbst κ -berührt, heißt κ -glatt in diesem Punkt. Ist es κ -glatt in jedem seiner Punkte, so nennen wir es κ -glatt schlechthin.

⁹⁾ Diese Ungleichung besagt, da ja $\kappa(p_1, p_m; p'_1, p'_m)$ existieren soll, unter anderem, daß $p_1 \neq p_m, p'_1 \neq p'_m$ ist. Für $\vartheta \leq \vartheta^*$ folgt dies übrigens auch aus Bedingung IV.

¹⁰⁾ Die Definition läßt sich leicht auch auf mehrfache Punkte ausdehnen, wovon wir aber der Bequemlichkeit halber hier absehen.

¹¹⁾ $p \xrightarrow{C} q$ bedeutet, daß p auf C „vor“ q kommt; d. h. es gibt Intervallwerte γ_1 und γ_2 , ($\alpha \leq \gamma_1 < \gamma_2 \leq \beta$), so daß p Bild von γ_1 und q Bild von γ_2 ist. Falls C kein Bogen ist, kann also gleichzeitig $p \xrightarrow{C} q$ und $q \xrightarrow{C} p$ sein.

¹²⁾ $C \cdot U(p_0, \varrho)$ bedeutet wie üblich den Durchschnitt von C mit der Umgebung $U(p_0, \varrho)$.

§ 7. Existenztheorem.

In einem kompakten, halbmtrischen Raum R mit gleichmäßig stetigen Abständen sei eine Funktion gemäß den Bedingungen von § 5 definiert.

In einem Punkt $p_0 \in R$ mögen zwei Zahlen $\varrho_0 > 0$ und $D_0 > 0$ existieren, so daß in der Umgebung $U(p_0, \varrho_0)$ folgendes gilt:

In jedem Punkt $p \in U(p_0, \varrho_0)$ beginnt ein Bogen¹³⁾ C_p , dessen Durchmesser $D(C_p) > D_0$ ist, und es gibt zu jedem $\varepsilon > 0$ ein $\varrho > 0$ von der Art, daß $U(p, \varrho) \subset U(p_0, \varrho_0)$ und

$$1 - \varepsilon \leq \kappa(q', r'; s'', t'') \leq 1 + \varepsilon$$

gilt, falls

$$\begin{aligned} q', r' &\in C_p \cdot U(p, \varrho); \quad s'', t'' \in C_{p''} \cdot U(p, \varrho) \\ (p' \in U(p, \varrho), \quad p'' \in U(p, \varrho), \quad q' \xrightarrow{C_{p'}} r', \quad s'' \xrightarrow{C_{p''}} t'') \end{aligned}$$

ist.

Dann gibt es einen in p_0 beginnenden Bogen C_0 — von endlicher oberer und positiver unterer Länge¹⁴⁾ —, der κ -glatt ist und in jedem seiner Punkte p die Kurve C_p κ -berührt.

Beweis.

§ 8. Vorbemerkungen.

1) Es bedeutet keine Beschränkung der Allgemeinheit, wenn wir annehmen, daß

$$1 - \vartheta^* \leq \kappa(q', r'; s'', t'') \leq 1 + \vartheta^*,$$

falls

$$\begin{aligned} q', r' &\in C_{p'} \cdot U(p_0, \varrho_0); \quad s'', t'' \in C_{p''} \cdot U(p_0, \varrho_0) \\ (p' \in U(p_0, \varrho_0), \quad p'' \in U(p_0, \varrho_0), \quad q' \xrightarrow{C_{p'}} r', \quad s'' \xrightarrow{C_{p''}} t'') \end{aligned}$$

ist und ϑ^* die nach der Bedingung IV von § 5 zu der Funktion κ existierende Zahl ist.

2) Daraus, daß es ein D_0 gibt, so daß $D(C_p) > D_0$ ist, folgt wegen der gleichmäßigen Stetigkeit der Abstände, daß es ein $d_0 > 0$ gibt, von der Art, daß zu jedem p und jedem $d \leq d_0$ mindestens ein $q \in C_p$ existiert, so daß $\delta(p, q) = d$ ist.

3) Wir bezeichnen mit n_0 die kleinste natürliche Zahl, für die

$$U\left(U\left(p_0, \frac{\varrho_0}{2}\right), \frac{d_0}{n_0}\right) \subset U\left(p_0, \frac{3}{4}\varrho_0\right)$$

ist.

¹³⁾ Daß C_p als Bogen und nicht als allgemeine Kurve angenommen wird, bedeutet, wie man sich leicht überzeugt, keine Beschränkung der Allgemeinheit.

¹⁴⁾ Man beweist übrigens leicht, daß für jede κ -glatte Kurve C

$$0 < \bar{\lambda}(C) \cdot \sqrt{1 - \Theta^*} \leq \underline{\lambda}(C) < \infty$$

ist, wobei Θ^* die Zahl aus Bedingung IV von § 5 ist.

Wegen der gleichmäßigen Stetigkeit der Abstände kann man offenbar immer ein solches n_0 finden.

4) Wir bezeichnen mit m_n diejenige natürliche Zahl, für die

$$1 + \frac{\varrho_0(n_0 + n)}{2d_0\sqrt{1+\Theta^*}} \leq m_n < 2 + \frac{\varrho_0 \cdot (n_0 + n)}{2d_0\sqrt{1+\Theta^*}}$$

ist, unter Θ^* die in der Bedingung IV von § 5 vorkommende, ebenso bezeichnete Zahl verstanden.

§ 9. Wir konstruieren nun folgendermaßen eine Polygonfolge:

$$P_n = \{p_{1,n}, \dots, p_{m_n,n}\}, \quad (n = 1, 2, \dots).$$

Für jedes n sei $p_{1,n} = p_0$. Für jedes i , ($i = 1, \dots, m_n - 1$) sei $p_{i+1,n}$ der erste auf $p_{i,n}$ folgende Punkt von $C_{p_{i,n}}$, für den $\delta(p_{i,n}, p_{i+1,n}) = \frac{d_0}{n_0 + n}$ ist.

Die Polygone P_n haben also folgende Eigenschaften. Es ist

$$\nu(P_n) = \frac{d_0}{n_0 + n} \quad \text{und} \quad \lambda(P_n) = \frac{m_n - 1}{n_0 + n} \cdot d_0.$$

Aus letzterem ergibt sich nach Vorbemerkung 4) von § 8

$$(7) \quad \frac{\varrho_0}{2} \cdot \frac{1}{\sqrt{1+\Theta^*}} \leq \lambda(P_n) < \frac{\varrho_0}{2} \cdot \frac{1}{\sqrt{1+\Theta^*}} + \frac{d_0}{n_0 + n}.$$

Wir zeigen nun, daß

$$C_{p_{i,n}} \big]_{p_{i,n}}^{p_{i+1,n}} \subset U(p_0, \tfrac{3}{4}\varrho_0) \quad (i = 1, \dots, m_n - 1).$$

Angenommen, dies sei nicht richtig, dann gäbe es eine kleinste Zahl k , ($1 \leq k \leq m_n - 1$), so daß auf $C_{p_{k,n}} \big]_{p_{k,n}}^{p_{k+1,n}}$ ein Punkt p^* mit $\delta(p_0, p^*) = \tfrac{3}{4}\varrho_0$ liegt.

Für $i < k$ ¹⁵⁾ gilt dagegen

$$C_{p_{i,n}} \big]_{p_{i,n}}^{p_{i+1,n}} \subset U(p_0, \tfrac{3}{4}\varrho_0),$$

also erst recht

$$p_{i,n}, p_{i+1,n} \subset U(p_0, \tfrac{3}{4}\varrho_0), \quad (i = 1, \dots, k - 1).$$

Nach Bemerkung 1) von § 8 haben wir daher

$$(8) \quad 1 - \Theta^* \leq \kappa(p_{i,n}, p_{i+1,n}; p_{j,n}, p_{j+1,n}) \leq 1 + \Theta^* \quad \left(\begin{matrix} i = 1, \dots, k - 1 \\ j = 1, \dots, k - 1 \end{matrix} \right).$$

¹⁵⁾ $C_{p_{i,n}} \big]_{p_{i,n}}^{p_{i+1,n}}$ bedeutet das Stück des Bogens $C_{p_{i,n}}$, das mit $p_{i,n}$ beginnt und mit $p_{i+1,n}$ endet.

¹⁶⁾ Hier wird $k > 1$ vorausgesetzt. Für $k = 1$ wäre $p^* \subset U(p_0, \frac{d_0}{n_0})$, also nach Vorbemerkung 3) von § 8 $\subset U(p_0, \frac{3}{4}\varrho_0)$, entgegen unserer Annahme.

Also haben wir nach der Bedingung IV von § 5, angewendet auf den Polygonabschnitt $\{p_{1,n}, \dots, p_{k,n}\}$,

$$\delta(p_{1,n}, p_{k,n}) \leq \sqrt{1 + \Theta^*} \cdot \sum_{i=1}^{k-1} \delta(p_{i,n}, p_{i+1,n})$$

und nach (7)

$$< \frac{\varrho_0}{2}.$$

Nun hat aber p^* , da zu $C_{p_{k,n}}$ gehörig, von $p_{k,n}$ einen Abstand $\leq \frac{d_0}{n_0 + n}$. Also ist

$$p^* \in U\left(p_{k,n}, \frac{d_0}{n_0}\right) \subset U\left(U\left(p_0, \frac{\varrho_0}{2}\right), \frac{d_0}{n_0}\right)$$

und nach Bemerkung 3) von § 8

$$\subset U(p_0, \frac{1}{2}\varrho_0),$$

im Widerspruch zu unserer Annahme, daß $\delta(p_0, p^*) = \frac{1}{2}\varrho_0$ sei.

(8) gilt also auch für $k = m_n$, so daß wir wegen der Bedingung IV von § 5 für jeden Abschnitt A von P_n (insbesondere für P_n selbst) finden

$$\lambda(A) \cdot \sqrt{1 - \Theta^*} \leq \mu(A) \leq \sqrt{1 + \Theta^*} \cdot \lambda(A).$$

Es sind also die Voraussetzungen der Sätze von § 3 und § 4 erfüllt, d. h. es existiert eine Häufungskurve C_0 der P_n , die natürlich in p_0 beginnt, ein Bogen ist und für deren obere bzw. untere Länge

$$\frac{\varrho_0}{2} \cdot \sqrt{\frac{1 - \Theta^*}{1 + \Theta^*}} \leq \underline{\lambda}(C_0) \leq \bar{\lambda}(C_0) \leq \frac{\varrho_0}{2}$$

gilt. Offenbar ist überdies $C_0 \subset U(p_0, \varrho_0)$.

§ 10. C_0 ist die gesuchte Kurve; d. h. C_0 ist κ -glatt und κ -berührt in jedem ihrer Punkte p die Kurve C_p .

Nach den Definitionen von § 6 müssen wir zeigen: Es gibt zu jedem $p \in C_0$ und jedem $\varepsilon > 0$ ein $\varrho > 0$, so daß

$$(9) \quad 1 - \varepsilon \leq \kappa(q, r; s, t) \leq 1 + \varepsilon,$$

$$(10) \quad 1 - \varepsilon \leq \kappa(q, r; s^*, t^*) \leq 1 + \varepsilon,$$

wie immer q, r, s, t unter folgenden Bedingungen gewählt seien:

$$(11) \quad q, r, s, t \in C_0 \cdot U(p, \varrho); \quad s^*, t^* \in C_p \cdot U(p, \varrho)$$

$$(q \xrightarrow{C_0} r, s \xrightarrow{C_0} t, s^* \xrightarrow{C_p} t^*).$$

Zu p und ε gibt es nach Bedingung III von § 5 Zahlen $\varrho_1 > 0$ und $\vartheta > 0$, so daß, wenn immer für zwei in der Umgebung $U(p, \varrho)$ liegende Polygone $P = \{p_1, \dots, p_m\}$, $P' = \{p'_1, \dots, p'_m\}$

$$(12) \quad 1 - \vartheta \leq \left[\begin{array}{l} \kappa(p_i, p_{i+1}; p_j, p_{j+1}) \\ \kappa(p'_k, p'_{k+1}; p'_l, p'_{l+1}) \\ \kappa(p_i, p_{i+1}; p'_l, p'_{l+1}) \end{array} \right] \leq 1 + \vartheta \quad \begin{array}{l} (i, j = 1, \dots, m-1) \\ (k, l = 1, \dots, m'-1) \end{array}$$

gilt, stets

$$(13) \quad 1 - \frac{\varepsilon}{2} \leq \kappa(p_1, p_m; p'_1, p'_m) \leq 1 + \frac{\varepsilon}{2}$$

ist.

Zu der in dieser Weise bestimmten Zahl ϑ gibt es nach der Voraussetzung des zu beweisenden Existenztheorems ein $\varrho_2 > 0$ von der Art, daß $U(p, \varrho_2) \subset U(p_0, \varrho_0)$ und

$$(14) \quad 1 - \vartheta \leq \kappa(q', r'; s'', t'') \leq 1 + \vartheta$$

ist, falls

$$(15) \quad \begin{aligned} q', r' &\subset C_{p'} \cdot U(p, \varrho_2); \quad s'', t'' \subset C_{p''} \cdot U(p, \varrho_2) \\ (p' &\subset U(p, \varrho_2), \quad p'' \subset U(p, \varrho_2), \quad q' \xrightarrow{C_{p'}} r', \quad s'' \xrightarrow{C_{p''}} t''). \end{aligned}$$

Wir setzen $\varrho_3 = \min(\varrho_1, \varrho_2)$. Wegen der gleichmäßigen Stetigkeit der Abstände im Raum R gibt es zu ϱ_3 ein $\varrho > 0$, so daß, wenn $q, r \subset C_0 \cdot U(p, \varrho)$ ist, stets $C_0]_q^r \subset U(p, \varrho_3)$ ist.

§ 11. ϱ ist bereits die gesuchte Zahl.

Seien nämlich sechs Punkte q, r, s, t, s^*, t^* gemäß Bedingung (11) fest gewählt.

Man kann wegen der gleichmäßigen Stetigkeit der Abstände eine Zahl $\Delta > 0$ finden, so daß

$$(16) \quad U(C_0]_q^r, \frac{\Delta}{2}) \subset U(p, \varrho_3),$$

$$(17) \quad U(C_0]_s^t, \frac{\Delta}{2}) \subset U(p, \varrho_3)$$

ist.

Außerdem sei, was nach der Bedingung II von § 5 möglich ist, Δ so angenommen, daß

$$(18) \quad |\kappa(q, r; s, t) - \kappa(\bar{q}, \bar{r}; \bar{s}, \bar{t})| < \frac{\varepsilon}{2},$$

$$(19) \quad |\kappa(q, r; s^*, t^*) - \kappa(\bar{q}, \bar{r}; s^*, t^*)| < \frac{\varepsilon}{2}$$

ist, falls

$$(20) \quad \delta(q, \bar{q}) + \delta(r, \bar{r}) < \Delta, \quad \delta(s, \bar{s}) + \delta(t, \bar{t}) < \Delta$$

ist.

Da C_0 Häufungskurve der im § 9 definierten Polygonfolge ist, enthält diese ein Polygon P_{n_j} , das einen Abschnitt

$$A_{n_j}^{(1)} = \{p_{k_1, n_j}, \dots, p_{l_1, n_j}, \dots, p_{l_1, n_j}\}$$

und einen Abschnitt

$$A_{n_j}^{(2)} = \{p_{k_2, n_j}, \dots, p_{l_2, n_j}, \dots, p_{l_2, n_j}\}$$

besitzt mit den folgenden Eigenschaften:

$$(21) \quad \delta(q, p_{k_1, n_J}) + \delta(r, p_{l_1, n_J}) < \Delta,$$

$$(22) \quad \delta(s, p_{k_2, n_J}) + \delta(t, p_{l_2, n_J}) < \Delta,$$

$$(23) \quad A_{n_J}^{(1)} \subset U(C_0]_{\eta}, \frac{\Delta}{2}), \quad A_{n_J}^{(2)} \subset U(C_0]_{\eta}, \frac{\Delta}{2}).$$

Wegen (16) und (17) folgt aus (23)

$$A_{n_J}^{(1)} \subset U(p, \varrho_3), \quad A_{n_J}^{(2)} \subset U(p, \varrho_3)$$

und nach Definition von ϱ_3 insbesondere

$$p_{i_1+1, n_J} \subset C_{p_{i_1, n_J}} U(p, \varrho_2), \quad p_{i_2+1, n_J} \subset C_{p_{i_2, n_J}} U(p, \varrho_2).$$

Setzen wir also

$$p' = q' = p_{i_1, n_J}; \quad p'' = s'' = p_{i_2, n_J}; \quad r' = p_{i_1+1, n_J}; \quad t'' = p_{i_2+1, n_J} \\ (i_1 = k_1, \dots, l_1 - 1; \quad i_2 = k_2, \dots, l_2 - 1),$$

so ist (15) erfüllt und es gilt daher (14), d. h. es ist

$$1 - \vartheta \leq \kappa(p_{i_1, n_J}, p_{i_1+1, n_J}; p_{i_2, n_J}, p_{i_2+1, n_J}) \leq 1 + \vartheta.$$

Entsprechend sieht man ein, daß auch

$$1 - \vartheta \leq \left\{ \begin{array}{l} \kappa(p_{i_1, n_J}, p_{i_1+1, n_J}; p_{j_1, n_J}, p_{j_1+1, n_J}) \\ \kappa(p_{i_2, n_J}, p_{i_2+1, n_J}; p_{j_2, n_J}, p_{j_2+1, n_J}) \\ \kappa(p_{i_1, n_J}, p_{i_1+1, n_J}; s^*, t^*) \end{array} \right\} \leq 1 + \vartheta \quad \begin{array}{l} (i_1, j_1 = k_1, \dots, l_1 - 1) \\ (i_2, j_2 = k_2, \dots, l_2 - 1) \end{array}$$

gilt.

Setzt man also in (12) für P den Abschnitt $A_{n_J}^{(1)}$ und für P' entweder $A_{n_J}^{(2)}$ oder $\{s^*, t^*\}$, so hat man nach (13)

$$(24) \quad 1 - \frac{\varepsilon}{2} \leq \kappa(p_{k_1, n_J}, p_{l_1, n_J}; p_{k_2, n_J}, p_{l_2, n_J}) \leq 1 + \frac{\varepsilon}{2},$$

$$(25) \quad 1 - \frac{\varepsilon}{2} \leq \kappa(p_{k_1, n_J}, p_{l_1, n_J}; s^*, t^*) \leq 1 + \frac{\varepsilon}{2}.$$

Da nun, wenn wir

$$\bar{q} = p_{k_1, n_J}, \quad \bar{r} = p_{l_1, n_J}, \\ \bar{s} = p_{k_2, n_J}, \quad \bar{t} = p_{l_2, n_J}$$

setzen, die Bedingungen (20) – wegen (21) und (22) – erfüllt sind, so gelten auch (18) und (19), so daß sich aus (24) und (25) die Behauptung [vgl. (9) und (10)] ergibt.

III. Die gewöhnlichen Differentialgleichungen und die Paratingentgleichungen¹⁷⁾ als Spezialfälle von κ -Problemen.

§ 12. Daß die κ -Probleme als eine Verallgemeinerung der gewöhnlichen Differentialgleichungen angesprochen werden können, sehen wir ein, indem wir, zu einem n -dimensionalen euklidischen Raum übergehend, $\kappa(p, q; p', q')$ als den Cosinus des Winkels der beiden durch die Punktepaare p, q und p', q' , ($p \neq q, p' \neq q'$), bestimmten, gerichteten Geraden definieren. Offenbar sind die Bedingungen I bis IV von § 5 erfüllt.

Bei dieser Wahl der Funktion κ nennen wir eine κ -glatte Kurve *glatt* schlechthin, und wenn sich zwei Kurven κ -berühren, sagen wir schlechthin, daß sie sich *berühren*, wobei dann die Begriffe „glatt“ und „sich berühren“ ersichtlich die übliche Bedeutung haben.

Das Existenztheorem von § 7 entspricht dann genau dem bekannten Satz, wonach durch jeden Punkt eines Gebietes G des n -dimensionalen Raumes mit den Koordinaten x_1, \dots, x_n , in dem die Funktionen $f_i(x_1, \dots, x_n)$, ($i = 1, \dots, n$) sämtlich überall stetig sind und $\sum_{j=1}^n f_j^2(x_1, \dots, x_n) > 0$ ist, mindestens eine Kurve $x_i(\tau)$ hindurch geht, für die $\frac{dx_i}{d\tau} = f_i(x_1, \dots, x_n)$ gilt¹⁸⁾.

§ 13. Es ist nicht uninteressant zu sehen, daß dieses κ -Problem auch in beliebigen halbmetrischen Räumen mit gleichmäßig stetigen Abständen seinen Sinn behält.

Bekanntlich läßt sich der Cosinus des Winkels zwischen den Strecken pq und $p'q'$, ($p \neq q, p' \neq q'$) folgendermaßen durch die Abstände zwischen den Punkten ausdrücken:

$$(26) \quad \cos(p, q; p', q') = \frac{-\delta(p, p')^2 - \delta(q, q')^2 + \delta(p, q')^2 + \delta(q, p')^2}{2 \delta(p, q) \cdot \delta(p', q')},$$

wobei δ den euklidischen Abstand bezeichnet.

Es liegt nahe, in einem beliebigen halbmetrischen Raum zu jedem Paar von Punktepaaren eine Zahl $\cos(p, q; p', q')$ gemäß (26) zu definieren, wobei jetzt δ den Abstand in diesem halbmetrischen Raum bedeutet.

Wir zeigen, daß $\cos(p, q; p', q')$ in einem halbmetrischen Raum mit gleichmäßig stetigen Abständen eine κ -Funktion ist; d. h. sie erfüllt die Bedingungen I bis IV von § 5. Die Bedingungen I, III, IV gelten sogar in einem beliebigen halbmetrischen Raum.

¹⁷⁾ Vgl. Fußnote 1).

¹⁸⁾ Um dies einzusehen, braucht man die in § 7 vorkommenden Kurven C_p in jedem Punkt $y = (x_1, \dots, x_n)$ von G nur als in p beginnende Strecke von der Länge 1 und den Richtungs-cosinus $\frac{f_i(x_1, \dots, x_n)}{\sum_{j=1}^n f_j^2(x_1, \dots, x_n)}$ zu wählen.

§ 14. Für die Bedingung I folgt dies trivialerweise aus (26).

Daß die Bedingung II erfüllt ist, erkennen wir leicht durch mehrfache Anwendung der Voraussetzung 2) von § 1, worauf wir der Kürze halber nicht eingehen wollen.

Zum Beweis der Bedingungen III und IV benötigen wir folgenden

Hilfssatz. Seien $P = \{p_1, \dots, p_m\}$ und $P' = \{p'_1, \dots, p'_{m'}\}$ zwei beliebige Polygone eines halbmetrischen Raumes, für die $p_1 \neq p_m$ und $p'_1 \neq p'_{m'}$ gelte und die den folgenden $(m-1) \cdot (m'-1)$ Bedingungen

$$(27) \quad \eta_2 \leq \cos(p_i, p_{i+1}; p'_j, p'_{j+1}) \leq \eta_1 \quad \left(\begin{matrix} i = 1, \dots, m-1 \\ j = 1, \dots, m'-1 \end{matrix} \right)$$

genügen, unter η_1, η_2 beliebige reelle Zahlen ($\eta_2 \leq \eta_1$) verstanden.

Dann ist

$$(28) \quad \eta_2 \cdot \lambda(P) \cdot \lambda(P') \leq \cos(p_1, p_m; p'_1, p'_{m'}) \cdot \mu(P) \cdot \mu(P') \leq \eta_1 \cdot \lambda(P) \cdot \lambda(P').^{19)}$$

Beweis. Die Ungleichungen (27) lassen sich, indem wir gleich summieren, folgendermaßen schreiben:

$$\begin{aligned} (29) \quad \eta_2 \cdot \sum_{\substack{i=1, \dots, m-1 \\ j=1, \dots, m'-1}} \delta(p_i, p_{i+1}) \cdot \delta(p'_j, p'_{j+1}) \\ \leq \frac{1}{2} \sum_{\substack{i=1, \dots, m-1 \\ j=1, \dots, m'-1}} [-\delta(p_i, p'_j)^2 - \delta(p_{i+1}, p'_{j+1})^2 + \delta(p_i, p'_{j+1})^2 + \delta(p_{i+1}, p'_j)^2] \\ \leq \eta_1 \cdot \sum_{\substack{i=1, \dots, m-1 \\ j=1, \dots, m'-1}} \delta(p_i, p_{i+1}) \cdot \delta(p'_j, p'_{j+1}). \end{aligned}$$

Nun gilt offenbar

$$(30) \quad \sum_{\substack{i=1, \dots, m-1 \\ j=1, \dots, m'-1}} \delta(p_i, p_{i+1}) \delta(p'_j, p'_{j+1}) = \lambda(P) \cdot \lambda(P').$$

Ferner hat man

$$\begin{aligned} (31) \quad \sum_{\substack{i=1, \dots, m-1 \\ j=1, \dots, m'-1}} [-\delta(p_i, p'_j)^2 - \delta(p_{i+1}, p'_{j+1})^2 + \delta(p_{i+1}, p'_j)^2 + \delta(p_i, p'_{j+1})^2] \\ = \sum_{i=1, \dots, m-1} [\delta(p_i, p'_m)^2 - \delta(p_i, p'_1)^2 + \delta(p_{i+1}, p'_1)^2 - \delta(p_{i+1}, p'_m)^2] \\ = \delta(p_1, p'_m)^2 - \delta(p_1, p'_1)^2 + \delta(p_m, p'_1)^2 - \delta(p_m, p'_m)^2. \end{aligned}$$

Setzt man (30) und (31) in (29), so erhält man (28).

¹⁹⁾ Falls $p_1 = p_m$ oder $p'_1 = p'_{m'}$, erhalten wir statt (28), wie aus dem Gang des Beweises sofort ersichtlich, $\eta_2 \leq 0 \leq \eta_1$.

Indem wir $P = P'$ setzen, erhalten wir folgendes *Korollar*.

Gentüge $P = \{p_1, \dots, p_m\}$ ²⁰⁾ den Bedingungen

$$1 - \vartheta_2 \leq \cos(p_i, p_{i+1}; p_j, p_{j+1}) \leq 1 + \vartheta_1 \quad \left(\begin{array}{l} 0 \leq \vartheta_2 < 1 \\ 0 \leq \vartheta_1 \end{array} \right).$$

Dann ist

$$\lambda(P) \cdot \sqrt{1 - \vartheta_2} \leq \mu(P) \leq \lambda(P) \cdot \sqrt{1 + \vartheta_1}.$$

Die Bedingung IV von § 5 ist also offensichtlich erfüllt. Das gleiche gilt aber auch von der Bedingung III. Durch Heranziehung des Hilfssatzes und des Korollars ergibt sich nämlich, daß, wenn für zwei Polygone $P = \{p_1, \dots, p_m\}$, $P' = \{p'_1, \dots, p'_m\}$ die Bedingungen

$$1 - \vartheta \leq \left\{ \begin{array}{l} \cos(p_i, p_{i+1}; p_j, p_{j+1}) \\ \cos(p'_k, p'_{k+1}; p'_l, p'_{l+1}) \\ \cos(p_i, p_{i+1}; p'_l, p'_{l+1}) \end{array} \right\} \leq 1 + \vartheta \quad (0 \leq \vartheta < 1)$$

gelten,

$$\frac{1 - \vartheta}{1 + \vartheta} \leq \cos(p_1, p_m; p'_1, p'_m) \leq \frac{1 + \vartheta}{1 - \vartheta}$$

ist.

Man braucht also, um Bedingung III zu erfüllen, nur zu jedem $\vartheta > 0$ ein $\vartheta > 0$ so zu wählen, daß

$$1 - \vartheta \leq \frac{1 - \vartheta}{1 + \vartheta}, \quad \frac{1 + \vartheta}{1 - \vartheta} \leq 1 + \vartheta$$

ist, was offenbar immer möglich ist. $\vartheta > 0$ kann beliebig angenommen werden.

§ 15. Wir beleuchten schließlich die Tragweite der Definition von § 5 noch dadurch, daß wir — nun wieder einen n -dimensionalen euklidischen Raum zugrunde legend — den folgenden Satz von Herrn Zaremba ²¹⁾ durch Wahl einer geeigneten κ -Funktion beweisen.

Satz. In jedem Punkt p eines gewissen Gebietes G eines n -dimensionalen euklidischen Raumes sei ein Büschel \mathfrak{B}_p von gerichteten Halbgeraden h mit folgenden Eigenschaften gegeben.

1) Die Gesamtheit aller, durch einen beliebigen Punkt gelegten Halbgeraden, die zu mindestens einer Halbgeraden eines der Büschel parallel sind, liegt auf einer Seite einer durch diesen Punkt gelegten $(n-1)$ -dimensionalen Ebene und bildet mit ihr einen Winkel vom Bogenmaß $\geq \varphi_0 > 0$.

2) Wenn zwei verschiedene Halbgeraden einem Büschel \mathfrak{B}_p angehören, so gehören zu ihm auch sämtliche dem kleineren durch die beiden Halbgeraden bestimmten Winkelraum angehörenden Halbgeraden.

²⁰⁾ Hier braucht, wie man unter Beachtung der Fußnote ¹⁹⁾ und der Ungleichung $\vartheta_1 < 1$ erkennt, $p_1 \neq p_m$ nicht eigens vorausgesetzt zu werden.

²¹⁾ Vgl. Zaremba, l. c. S. 149.

3) Zu jedem Punkt $p \in G$ und jedem $\varepsilon > 0$ gibt es ein $\varrho > 0$, so daß, wenn $h' \in \mathfrak{B}_p$, ($p' \in U(p, \varrho)$) ist, es im Büschel \mathfrak{B}_p ein h gibt, so daß h und h' einen Winkel vom Bogenmaß $< \varepsilon$ bildet.

Dann beginnt in jedem Punkt $p_0 \in G$ ein Bogen C_0 , der in jedem seiner Punkte p folgende Eigenschaft besitzt:

Es gibt zu jedem $\varepsilon > 0$ ein $\varrho^* > 0$, so daß, wenn $p', p'' \in C_0 \cdot U(p, \varrho^*)$, ($p' \xrightarrow{C_0} p''$) gilt, es in \mathfrak{B}_p eine Halbgerade gibt, die mit der durch p', p'' bestimmten einen Winkel $< \varepsilon$ bildet.

§ 16. Beweis. Wir definieren eine Funktion $\text{par}(p, q; p', q')$, ($p \neq q$, $p' \neq q'$), in der folgenden Weise:

Es sei $\text{par}(p, q; p', q')$ die obere Grenze aller Zahlen $1 - \alpha$, ($\alpha > 0$) von der Art, daß es einen Punkt p_0 gibt, so daß 1) $p, p' \in U(p_0, \alpha)$ gilt und 2) in \mathfrak{B}_{p_0} ein h und ein h' ²²⁾ vorhanden sind, so daß h mit der durch p, q und h' mit der durch p', q' bestimmten Halbgeraden einen Winkel vom Bogenmaß $\leq \alpha$ bildet.

Wie wir in den folgenden Paragraphen zeigen, ist $\text{par}(p, q; p', q')$ eine κ -Funktion.

Nach dem Existenztheorem von § 7 beginnt daher in jedem Punkt p_0 des Gebietes G ein Bogen C_0 , der in jedem seiner Punkte p mindestens eine der Halbgeraden $\subset \mathfrak{B}_p$ — im Sinne der Funktion $\text{par}(p, q; p', q')$ — κ -berührt.

C_0 ist bereits das gesuchte Streckenbild.

Zunächst gibt es nämlich nach Voraussetzung 3) des zu beweisenden Satzes zu jedem $\frac{\varepsilon}{2} > 0$ und jedem Punkt $p \in C_0$ ein $\varrho > 0$, so daß, wenn $p^* \in U(p, \varrho)$ ist, zu jeder Halbgeraden $h^* \in \mathfrak{B}_p$ eine Halbgerade $\bar{h} \in \mathfrak{B}_p$ existiert, die mit h^* einen Winkel $< \frac{\varepsilon}{2}$ bildet. ϱ sei noch $< \varepsilon$ gewählt.

Zu $\frac{\varrho}{2} > 0$ gibt es nun, da ja C_0 in p mindestens eine der Halbgeraden von \mathfrak{B}_p κ -berührt, nach § 6 und nach der Definition von $\text{par}(p, q; p', q')$ ein $\varrho_1 > 0$, so daß, wenn $p', p'' \in C_0 \cdot U(p, \varrho_1)$, ($p' \xrightarrow{C_0} p''$) ist, es ein $p^* \in U(p, \frac{\varrho}{2}) \subset U(p, \varrho_1 + \frac{\varrho}{2})$ gibt, so daß in \mathfrak{B}_p mindestens eine Halbgerade vorhanden ist, die mit der durch p', p'' bestimmten einen Winkel $< \frac{\varrho}{2}$, also $< \frac{\varepsilon}{2}$ bildet. Es genügt daher $\varrho^* = \min(\varrho_1, \frac{\varrho}{2})$ zu wählen, womit der Satz von § 15 bewiesen ist.

§ 17. Es muß lediglich nachträglich gezeigt werden, daß $\text{par}(p, q; p', q')$ tatsächlich die Bedingungen I bis IV von § 5 erfüllt.

Was die Bedingung I betrifft, so ist dies selbstverständlich.

²²⁾ h und h' können auch zusammenfallen.

Bedingung II. Wir zeigen:

Zu jedem Punktepaar p, q , ($p \neq q$), und jedem $\varepsilon > 0$ gibt es ein $\Delta > 0$ ²³⁾, so daß aus

$$(32) \quad \delta(p, \bar{p}) + \delta(q, \bar{q}) < \Delta$$

folgt

$$(33) \quad |\text{par}(p, q; p', q') - \text{par}(\bar{p}, \bar{q}; p', q')| < \varepsilon.$$

In der Tat kann man offenbar ein Δ finden, so daß $0 < \Delta < \frac{\varepsilon}{2}$ ist und aus (32) stets folgt, daß die durch p, q bzw. \bar{p}, \bar{q} bestimmten Halbgeraden einen Winkel $< \frac{\varepsilon}{2}$ bilden.

Sei nun $\text{par}(\bar{p}, \bar{q}; p', q') > 1 - \alpha$, ($\alpha > 0$). Dann gibt es einen Punkt p_0 , so daß $\bar{p}, \bar{q} \subset U(p_0, \alpha)$ und in \mathfrak{B}_{p_0} Halbgeraden enthalten sind, die mit \bar{p}, \bar{q} bzw. p', q' Winkel $< \alpha$ bilden.

Dann ist aber wegen (32)

$$p \subset U(p_0, \alpha + \Delta) \subset U(p_0, \alpha + \frac{\varepsilon}{2})$$

und \bar{p} bildet mit der durch p, q bestimmten Halbgeraden einen Winkel $< \alpha + \frac{\varepsilon}{2}$. Also ist

$$\text{par}(p, q; p', q') > 1 - \alpha - \frac{\varepsilon}{2}.$$

Nach Definition der Funktion par kann schließlich α so gewählt werden, daß

$$\text{par}(p, q; p', q') > \text{par}(\bar{p}, \bar{q}; p', q') - \varepsilon.$$

Entsprechend beweist man

$$\text{par}(\bar{p}, \bar{q}; p', q') > \text{par}(p, q; p', q') - \varepsilon,$$

womit (33) bewiesen ist.

Bedingung IV. Wir zeigen, daß für jedes Polygon $P = \{p_1, \dots, p_m\}$, für das

$$(34) \quad \text{par}(p_i, p_{i+1}; p_j, p_{j+1}) \geq 1 - \frac{\varphi_0}{2} \quad \begin{matrix} (i = 1, \dots, m-1) \\ (j = 1, \dots, m-1) \end{matrix}$$

ist, stets

$$(35) \quad \frac{\mu(P)}{\lambda(P)} \geq \sin \frac{\varphi_0}{2}$$

gilt.

In der Tat folgt aus (34) nach der Definition der Funktion par , daß es, wie auch $\varepsilon > 0$ gewählt sei, zu jedem Punktepaar p_i, p_{i+1} , ($i = 1, \dots, m-1$)

²³⁾ p', q' , ($p' \neq q'$), können beliebig gewählt werden.

²⁴⁾ φ_0 ist die in der Voraussetzung 1) des Satzes von § 15 vorkommende Zahl.

einen Punkt $p \in G$ gibt, so daß ein $h \in \mathfrak{B}_p$ existiert von der Art, daß der Winkel zwischen h und $h_{p_i, p_{i+1}}$ ²⁵⁾

$$\angle(h, h_{p_i, p_{i+1}}) < \frac{\varphi_0}{2} + \varepsilon$$

ist.

Nach der Voraussetzung 1) des Satzes von § 15 folgt, daß es durch einen beliebigen Punkt des Raumes zu den $h_{p_i, p_{i+1}}$ ($i = 1, \dots, m-1$) parallele Halbgeraden gibt, die alle auf einer Seite einer $(n-1)$ -dimensionalen Ebene liegen und mit ihr einen Winkel $\geq \frac{\varphi_0}{2}$ bilden. Dann aber gilt offenbar (35).

§ 18. Bedingung III. Zum Beweis, daß auch diese Bedingung erfüllt ist, benötigen wir den folgenden leicht zu beweisenden

Hilfssatz. Zu jedem a , ($0 < a < 1$) und jedem $\Theta > 0$ gibt es ein $f(\Theta, a) > 0$, so daß, wenn für irgend zwei aus beliebig (aber gleich) vielen Punkten bestehende Polygone $P = \{p_1, \dots, p_m\}$, $\bar{P} = \{\bar{p}_1, \dots, \bar{p}_m\}$ eines euklidischen Raumes

$$\frac{\mu(P)}{\lambda(P)} \geq a; \quad d(p_i, p_{i+1}) = d(\bar{p}_i, \bar{p}_{i+1}) \quad (i = 1, \dots, m),$$

$$\angle(p_i, p_{i+1}; \bar{p}_i, \bar{p}_{i+1}) < f(\Theta, a)$$

gilt, stets

$$\angle(p_1, p_m; \bar{p}_1, \bar{p}_m) < \Theta$$

ist.

Insbesondere setzen wir im folgenden

$$(36) \quad f\left(\Theta, \sin \frac{\varphi_0}{2}\right) = \vartheta_1,$$

wobei φ_0 die in Voraussetzung 1) des Satzes von § 15 vorkommende Zahl bedeute.

Wir zeigen nun, daß man zu jedem $\Theta > 0$ und jedem Punkt $p \in G$ ein $\vartheta > 0$ und $\varrho > 0$ so bestimmen kann, daß, wenn immer für zwei Polygone $P = \{p_1, \dots, p_m\}$ und $P' = \{p'_1, \dots, p'_{m'}\}$, die in $U(p, \varrho)$ liegen,

$$(37) \quad 1 - \vartheta \leq \begin{cases} \text{par}(p_i, p_{i+1}; p_j, p_{j+1}), & (i, j = 1, \dots, m-1) \\ \text{par}(p'_k, p'_{k+1}; p'_l, p'_{l+1}) & (k, l = 1, \dots, m'-1) \end{cases}$$

gilt,

$$(38) \quad p_1, p'_1 \subset U(p, \Theta)$$

ist, und eine Halbgerade h bzw. $h' \in \mathfrak{B}_p$ existiert, so daß

$$(39) \quad \angle(h; p_1, p_m) \leq \Theta, \quad \angle(h'; p'_1, p'_{m'}) \leq \Theta$$

ist. Daß dann die Bedingung III von § 5 erst recht erfüllt ist, ist klar.

²⁵⁾ $h_{p_i, p_{i+1}}$ bedeutet die durch p_i, p_{i+1} bestimmte, von p_i nach p_{i+1} zielende Halbgerade.

Zunächst kann man zu jedem $p \in G$ ein $\varrho_1 > 0$ gemäß Voraussetzung 3) des Satzes von § 15 wählen, so daß zu jedem $h' \in \mathfrak{B}_{p'}$, ($p' \in U(p, \varrho_1)$) ein $h \in \mathfrak{B}_p$ existiert, das die Ungleichung $\angle(h, h') < \frac{\vartheta_1}{2}$ erfüllt.

Wir beweisen nun, daß es genügt

$$\varrho < \min\left(\Theta, \frac{\varrho_1}{2}\right); \quad \vartheta < \min\left(\frac{\vartheta_1}{4}, \frac{\vartheta_0}{2}, \frac{\vartheta_1}{4}\right)$$

zu wählen.

Seien $P = \{p_1, \dots, p_m\}$ und $P' = \{p'_1, \dots, p'_{m'}\}$ zwei beliebige, aber im folgenden fest gewählte Polygone $\subset U(p, \varrho)$, die die Ungleichungen (37) erfüllen.

Aus $\varrho < \Theta$ folgt (38).

Aus (37) folgt, daß es Punkte $p_{i,j}, p'_{k,l}$, ($i, j = 1, \dots, m-1$; $k, l = 1, \dots, m'-1$) gibt, von der Art, daß

$$(40) \quad p_i, p_j \in U(p_{i,j}, 2\vartheta); \quad p'_k, p'_l \in U(p'_{k,l}, 2\vartheta)$$

und Halbgeraden $h_{p_i}, h_{p_j} \in \mathfrak{B}_{p_{i,j}}$ und $h'_{p'_k}, h'_{p'_l} \in \mathfrak{B}_{p'_{k,l}}$ existieren, so daß

$$(41) \quad \left. \begin{aligned} \angle(h_{p_i}; p_i, p_{i+1}), \\ \angle(h_{p_j}; p_j, p_{j+1}) \end{aligned} \right\} \leq 2\vartheta; \quad \left. \begin{aligned} \angle(h'_{p'_k}; p'_k, p'_{k+1}), \\ \angle(h'_{p'_l}; p'_l, p'_{l+1}) \end{aligned} \right\} \leq 2\vartheta$$

ist.

Da $\varrho < \frac{\varrho_1}{2}$ und $\vartheta < \frac{\vartheta_1}{4}$ angenommen wurde, folgt aus (40)

$$p_{i,j}, p'_{k,l} \in U(p, \varrho_1).$$

Wegen der Definition von ϱ_1 ergibt sich daher aus (41), daß \mathfrak{B}_p Halbgeraden h_i und h'_k enthält, so daß

$$\left. \begin{aligned} \angle(h_i; p_i, p_{i+1}), \\ \angle(h'_k; p'_k, p'_{k+1}) \end{aligned} \right\} < 2\vartheta + \frac{\vartheta_1}{2}$$

ist, und wegen

$$\vartheta < \frac{\vartheta_1}{4}$$

$$(42) \quad < \vartheta_1.$$

Wir betrachten nun zwei Polygone $\bar{P} = \{\bar{p}_1, \dots, \bar{p}_m\}$ und $\bar{P}' = \{\bar{p}'_1, \dots, \bar{p}'_{m'}\}$ von folgender Art:

Es seien

$$(43) \quad \begin{aligned} d(p_i, p_{i+1}) &= d(\bar{p}_i, \bar{p}_{i+1}), & (i = 1, \dots, m-1) \\ \angle(h_i; \bar{p}_i, \bar{p}_{i+1}) &= 0 \end{aligned}$$

und

$$(44) \quad \begin{aligned} d(p'_k, p'_{k+1}) &= d(\bar{p}'_k, \bar{p}'_{k+1}), & (k = 1, \dots, m'-1) \\ \angle(h'_k; \bar{p}'_k, \bar{p}'_{k+1}) &= 0. \end{aligned}$$

Aus (43) und (44) folgt unter Berücksichtigung der Voraussetzung 2) des Satzes von § 15, daß es ein \bar{h} und $\bar{h}' \subset \mathfrak{B}_p$ gibt, so daß

$$(45) \quad \Delta(\bar{h}; \bar{p}_1, \bar{p}_m) = 0; \quad \Delta(\bar{h}'; \bar{p}'_1, \bar{p}'_m) = 0$$

ist.

Wenden wir nun den am Anfang dieses Paragraphen aufgeführten Hilfsatz auf P und \bar{P} bzw. P' und \bar{P}' an, wobei wir beachten, daß wegen (34), (35) und (37) und wegen $\vartheta < \frac{\tau_0}{2}$

$$\frac{\mu(P)}{\lambda(P)} \geq \sin \frac{\tau_0}{2}; \quad \frac{\mu(P')}{\lambda(P')} \geq \sin \frac{\tau_0}{2}$$

ist, so folgt aus (42) und (36)

$$\Delta(p_1, p_m, \bar{p}_1, \bar{p}_m) < \Theta; \quad \Delta(p'_1, p'_m, \bar{p}'_1, \bar{p}'_m) < \Theta,$$

woraus sich wegen (45) die Behauptung (39) ergibt.

(Eingegangen am 1. 4. 1942.)

Theorie der Poincaréschen Reihen zu den hyperbolischen Fixpunktsystemen der Hilbertschen Modulgruppe.

Von

Hans Maaß in Heidelberg.

Die Frage nach der Darstellbarkeit automorpher Formen zur Hilbertschen Modulgruppe durch explizite analytische Ausdrücke kann insofern als gelöst betrachtet werden, als bewiesen wurde¹⁾, daß die Poincaréschen Reihen zu den parabolischen Fixpunkten bereits alle automorphen Formen von der Dimension $-r < -2$ zu einem Multiplikatorsystem v vom Betrag Eins darstellen. Der etwas komplizierte Bau dieser Poincaréschen Reihen ist dadurch bedingt, daß jeder parabolische Fixpunkt der Hilbertschen Modulgruppe (im Fall der Variablenzahl $n > 1$) zugleich hyperbolischer Fixpunkt ist. Hält man Umschau nach einfacheren Grundelementen für die Theorie der automorphen Formen, so wird man zunächst darauf geführt, die zahlreichen Typen Poincaréscher Reihen systematisch auf ihre Leistungsfähigkeit zu untersuchen. Es ist Aufgabe des vorliegenden Aufsatzes, für zwei bisher noch nicht betrachtete Typen die analoge Theorie zu entwickeln, wie sie für die Reihen zu den parabolischen Fixpunkten bereits bekannt ist¹⁾. Es handelt sich dabei erstens um die Poincaréschen Reihen zu den hyperbolischen Fixpunktsystemen, zweitens um die Reihen zu den inneren Punkten des Bereichs \mathfrak{H} , definiert durch

$$(1) \quad \Im \tau^{(v)} > 0 \quad \text{für} \quad v = 1, 2, \dots, n.$$

Die Zuordnung von Poincaréschen Reihen zu Fixpunktsystemen bzw. Punkten ist im Sinne des von H. Petersson beschriebenen Prinzips der Erzeugung Poincaréscher Reihen zu verstehen²⁾. Im Falle einer Veränderlichen ($n = 1$) ist die Theorie der beiden genannten Typen von H. Petersson bereits vollständig entwickelt und niedergelegt²⁾. Die Art des Vorgehens bewährt sich auch im Fall mehrerer Veränderlicher ($n > 1$). Es zeigt sich jedoch, daß in diesem Fall die Mannigfaltigkeit der Poincaréschen Reihen zu den hyperbolischen Fixpunktsystemen, die frei von parabolischen Fixpunkten sind, wesentlich größer ist als im Fall $n = 1$. Der Grund liegt darin, daß nicht alle

¹⁾ H. Maaß, Zur Theorie der automorphen Funktionen von n Veränderlichen. Math. Annalen 117 (1940), S. 538–578.

²⁾ H. Petersson, Einheitliche Begründung der Vollständigkeitsätze für die Poincaréschen Reihen von reeller Dimension bei beliebigen Grenzkreisgruppen von erster Art. Abhandl. aus d. Math. Seminar d. Hansischen Univ. 14 (1941), S. 22–60.

Konjugierten $S^{(v)}$, $v = 1, 2, \dots, n$, einer hyperbolischen Substitution S aus der Hilbertschen Modulgruppe selbst hyperbolisch zu sein brauchen. Kommen genau t hyperbolische Konjugierte von S vor und sind dies etwa die t ersten, so sind die restlichen $(n - t)$ Konjugierten elliptisch. Für t sind alle Werte der Reihe $1, 2, \dots, n$ möglich. Betrachtet man die Poincaréschen Reihen zu dem Fixpunktsystem von S , so liefert $t = n$ das Analogon zum Peterssonschen \mathcal{E} -Typus²⁾, während die Reihen zu $1 \leq t < n$ eine Mischung von \mathcal{E} - und \mathcal{P} -Typus²⁾ darstellen. In den t ersten Veränderlichen haben sie nämlich den Charakter einer \mathcal{E} -Reihe, in den $(n - t)$ letzten Veränderlichen dagegen den Charakter einer \mathcal{P} -Reihe. $t = 0$ bedeute, daß an Stelle des Fixpunktsystems einer hyperbolischen Substitution ein beliebiger innerer Punkt von \mathfrak{H} tritt. Die zugehörigen Poincaréschen Reihen stellen den zweiten zu untersuchenden Typus dar und sind das Analogon zum \mathcal{P} -Typus. Im Verlauf der Untersuchung wird ein Formalismus entwickelt, der eine einheitliche Behandlung aller Fälle $0 \leq t \leq n$ gestattet. Was die Allgemeinheit der Ansätze anbelangt, so sind sie natürlich für eine beliebige diskontinuierliche hyperabelsche Gruppe durchführbar, sofern nur für diese Gruppen bezüglich ihrer hyperbolischen Fixpunktsysteme ähnliche Vollständigkeitsannahmen gemacht werden, wie sie bei der Definition der parabolischen Spitzen notwendig sind³⁾. Um aber klare Voraussetzungen zu schaffen, wollen wir uns auf die Hilbertsche Modulgruppe und ihre Untergruppen von endlichem Index beschränken. Mit \mathbf{G} soll, sofern nicht ausdrücklich etwas anderes bemerkt wird, eine solche Untergruppe bezeichnet werden. Die Untersuchung liefert im einzelnen folgende Resultate:

Es sei s für $t > 0$ das System derjenigen Fixpunkte einer hyperbolischen Transformation $S \in \mathbf{G}$, welche auf dem Rand von \mathfrak{H} liegen, s enthalte keine parabolischen Fixpunkte und unter den Konjugierten von S mögen genau t selbst hyperbolisch sein (etwa die t ersten). Die Punkte $\tau \in s$ sind dann von der Gestalt

$$(2) \quad \begin{aligned} \tau^{(v)} &= \tau_0^{(v)} \text{ oder } \tau_0^{(v)'} \text{ (reell) für } v = 1, 2, \dots, t, \\ \tau^{(v)} &= \tau_0^{(v)} \text{ (mit pos. Imaginärteil) für } v = t + 1, t + 2, \dots, n. \end{aligned}$$

Im Fall $t = 0$ bestehe s aus einem einzigen inneren Punkt τ_0 von \mathfrak{H} . Die Transformation

$$A = \begin{pmatrix} a_1 & a_2 \\ a_1' & a_2' \end{pmatrix}$$

habe die Eigenschaft, den Bereich $\Im m \tau^{(v)} > 0$ für $v \leq t$ in sich, für $v > t$ in den Einheitskreis und s in die spezielle Lage:

$$A\tau_0 = 0, \quad A\tau_0' = \infty$$

³⁾ H. Maaß, Über Gruppen von hyperabelschen Transformationen. Sitzungsberichte der Heidelberger Akademie der Wissenschaften, math.-naturwiss. Klasse (1940), 2. Abhandlung.

überzuführen; dabei bedeutet noch $\tau_0^{(r)} = \overline{\tau_0^{(r)}}$ für $r > t$. Jede (ganze) automorphe Form $f(\tau) \in \{G, -r, v\}$ besitzt eine Entwicklung nach gewissen „Ortsuniformisierenden“ $\omega_1, \omega_2, \dots, \omega_n$ des Punktsystems s . Sie hat die Gestalt

$$(3) \quad h(\tau) f(\tau) = \sum b_{k_1 k_2 \dots k_n} \omega_1^{k_1 + x_1} \dots \omega_t^{k_t + x_t} \omega_{t+1}^{k_{t+1}} \dots \omega_n^{k_n},$$

wobei über alle ganzzahligen k_μ mit

$$-\infty < k_\mu < \infty \text{ für } \mu \leq t \text{ und } 0 \leq k_\mu < \infty \text{ für } \mu > t$$

summiert wird. In der Entwicklung (3) ist zur Abkürzung

$$(4) \quad \begin{aligned} h(\tau) &= \prod_{r=1}^t (a_1^{(r)} \tau^{(r)} + a_2^{(r)})^2 (a_1^{(r)} \tau^{(r)} + a_2^{(r)})^2 \prod_{r=t+1}^n (a_1^{(r)} \tau^{(r)} + a_2^{(r)})^2, \\ \omega_\mu &= \prod_{r=1}^t (A^{(r)} \tau^{(r)})^{2\pi i m_\mu^{(r)}} \quad \text{für } \mu \leq t, \\ \omega_\mu &= A^{(u)} \tau^{(u)} \prod_{r=1}^t (A^{(r)} \tau^{(r)})^{2\pi i m_\mu^{(r)}} \quad \text{für } \mu > t \end{aligned}$$

gesetzt. $m_\mu^{(r)}$ und x_μ sind reelle Konstanten, die nur von G, A, v und einer Basis für die Gruppe $A(s)$ der Substitutionen aus G , welche s festlassen, abhängen. Durch geeignete Auswahl eines Zweiges der komplexen Logarithmenfunktion wird für die Fixierung der auftretenden vieldeutigen Funktionen gesorgt. Die Reihenentwicklung (3) besitzt in ganz \mathfrak{H} Gültigkeit. Die Funktionalgleichung

$$f(\tau) = \bar{v}(L) N(\gamma\tau + \delta)^{-r} f(L\tau) \quad (v\bar{v} = 1)$$

für eine Substitution $L \in G$ mit der zweiten Zeile $L = (\gamma, \delta)$ liefert zu jedem $L \in G$ eine neue Entwicklung wie folgt. In der Darstellung, die man für $f(\tau)$ aus (3) nach Division durch $h(\tau)$ erhält, wird τ durch $L\tau$ ersetzt. Das Resultat geht nach Multiplikation mit $\bar{v}(L) N(\gamma\tau + \delta)^{-r}$ wieder in $f(\tau)$ über. \mathfrak{S} sei ein volles System von Substitutionen aus G , für welche auf diese Weise wesentlich verschiedene Entwicklungen entstehen. Setzt man in der Entwicklung zu L formal einen Koeffizienten, etwa $b_{k_1 k_2 \dots k_n}$, gleich 1, alle anderen Koeffizienten gleich 0 und summiert über alle $L \in \mathfrak{S}$, so entsteht eine Poincarésche Reihe $\Xi_{-,r}(\tau) = \Xi_{-,r}(\tau, v, A, G, (k))$. Sie ist für $r > 2$, $|v| = 1$ in ganz \mathfrak{H} konvergent und stellt für jedes mögliche Exponentensystem (k) eine ganze Spitzenform vom Typus $\{G, -r, v\}$ dar. Nunmehr wird auf die Ξ -Reihen die Metrisierungstheorie für die automorphen Formen angewendet. Die Durchführung einer etwas längeren Rechnung ergibt für das Skalarprodukt einer Spitzenform $f(\tau) \in \{G, -r, v\}$ mit der Poincaréschen Reihe $\Xi_{-,r}(\tau)$ den Wert

$$(5) \quad f(\tau), \Xi_{-,r}(\tau, v, A, G, (k)) = C b_{k_1 k_2 \dots k_n}$$

wobei $b_{k_1 k_2 \dots k_n}$ der Entwicklungskoeffizient von $f(\tau)$ zum Exponentensystem k_1, k_2, \dots, k_n und $C = C(\mathbf{G}, \tau, v, A, (k))$ eine elementare Konstante, die sich im wesentlichen aus Werten der Γ -Funktion zusammensetzt. Diese sogenannte Grundformel ist die Quelle aller weiteren Sätze. Aus ihr ergeben sich die Kennzeichnung aller Poincaréschen Reihen $\Xi_{-,\tau}(\tau)$ durch Orthogonalitätsrelationen und der Beweis des Vollständigkeitssatzes, der besagt, daß unter den Poincaréschen Reihen $\Xi_{-,\tau}(\tau, v, A, \mathbf{G}, (k))$ zu festem A eine Basis für die Schar der Spitzenformen vom Typus $\{\mathbf{G}, -\tau, v\}$ aufgefunden werden kann.

§ 1.

Entwicklungen automorpher Formen zu hyperbolischen Fixpunktsystemen.

Es sei K ein total-reeller algebraischer Zahlkörper vom Absolutgrad n , \mathbf{M} die (engere) Hilbertsche Modulgruppe zu K und \mathbf{G} eine Untergruppe aus \mathbf{M} von endlichem Index. Eine hyperbolische Substitution

$$S = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathbf{G}$$

denken wir uns so ausgewählt, daß unter den Fixpunkten von S keine parabolischen vorkommen. Das ist genau dann der Fall, wenn $\gamma \neq 0$ und $\Delta = (\alpha + \delta)^2 - 4$ kein Quadrat einer Zahl aus K ist. Die 2ⁿ Fixpunkte τ von S bestimmt man durch die beiden Punkte

$$(6) \quad \begin{aligned} \tau_0 &= \frac{\alpha - \delta}{2\gamma} + \frac{1}{2|\gamma|} \sqrt{\Delta} \\ \tau'_0 &= \frac{\alpha - \delta}{2\gamma} - \frac{1}{2|\gamma|} \sqrt{\Delta} \end{aligned} \quad (\sqrt{\Delta} > 0 \text{ oder } -i\sqrt{\Delta} > 0),$$

indem man $\tau^{(v)} = \tau_0^{(v)}$ oder $\tau^{(v)'} = \tau'_0$ für $v = 1, 2, \dots, n$ unabhängig voneinander setzt. Es soll nun angenommen werden, daß etwa

$$\Delta^{(v)} > 0 \text{ für } v \leq t \quad \text{und} \quad \Delta^{(v)} < 0 \text{ für } v > t \quad (t \geq 1).$$

Infolgedessen sind $\tau_0^{(v)}$ und $\tau_0^{(v)'}$ für $v < t$ reell, für $v > t$ dagegen konjugiert komplex. Die 2^t Fixpunkte von S , die auf dem Rand von \mathfrak{H} liegen, werden zu einem System s vereinigt. Durch das Fixpunktsystem s ist auf Grund von (6) eine quadratische Erweiterung $\Omega = K(\tau_0 - \tau_0')$ über K eindeutig bestimmt. Eine beliebige Zahl $\omega \in \Omega$ werde durch den erzeugenden Automorphismus von Ω/K in ω' übergeführt. Diese Festsetzung steht im Einklang mit der Zuordnung: $\tau_0 \rightarrow \tau_0'$. Ausgehend von s bestimmen wir nun die abelsche Gruppe $\mathbf{A}(s)$ aller Substitutionen aus \mathbf{G} , welche s festlassen, und beweisen, daß $\mathbf{A}(s)$ t Substitutionen mit unabhängigen Multiplikatoren enthält. Multiplikatoren λ^2 sollen unabhängig heißen, wenn die zugehörigen Vektoren $\{\log |\lambda^{(1)}|^2, \log |\lambda^{(2)}|^2, \dots, \log |\lambda^{(n)}|^2\}$ in bezug auf den Körper der

reellen Zahlen linear unabhängig sind. Wie man leicht sieht, gibt es höchstens t Substitutionen in $\mathbf{A}(s)$ mit unabhängigen Multiplikatoren. Um die erste der Behauptungen einzusehen, beachten wir, daß jede Substitution $L \in \mathbf{A}(s)$ in der Gestalt

$$(7) \quad L = Q^{-1} \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} Q \quad \text{mit} \quad Q = \begin{pmatrix} 1 & -\tau_0 \\ 1 & -\tau_0 \end{pmatrix}$$

angesetzt werden kann; dabei ist λ^2 der Multiplikator von L . Sind α und δ die Diagonalelemente von L , so ist offenbar $\alpha + \delta = \lambda + \lambda^{-1}$ oder

$$\lambda^2 - (\alpha + \delta)\lambda + 1 = 0,$$

d. h. λ eine Einheit aus Ω mit Relativnorm 1:

$$(8) \quad \lambda\lambda' = 1.$$

Es genügt, den beabsichtigten Beweis für die Hilbertsche Modulgruppe M an Stelle von G zu führen; denn eine geeignet hohe Potenz einer jeden Substitution aus M liegt ja auch in G . Wir müssen also Einheiten $\lambda \in \Omega$ mit Relativnorm 1 derart bestimmen, daß die Koeffizienten von L nicht nur, wie aus der Darstellung (7) leicht zu ersehen ist, in K liegen, sondern überdies ganze Zahlen sind. Das ist bestimmt dann der Fall, wenn für ein gewisses auftretendes ganzes Nennerideal $\mathfrak{a} \subset \Omega$ die Kongruenz

$$(9) \quad \lambda \equiv 1 \pmod{\mathfrak{a}}$$

befriedigt wird. Die Frage nach Substitutionen mit unabhängigen Multiplikatoren läuft also darauf hinaus, unabhängige Einheiten in Ω zu finden, welche den Bedingungen (8) und (9) genügen. Nach Dirichlet ist die genaue Anzahl der Grundeinheiten in Ω gleich $n - 1 + t$, also um t größer als die Anzahl der Grundeinheiten in K . Elementare Überlegungen liefern dann t unabhängige Einheiten λ mit den Eigenschaften (8) und (9). Damit ist nun die Behauptung über $\mathbf{A}(s) \subset G$ bewiesen. $\lambda_1^2, \lambda_2^2, \dots, \lambda_t^2$ sei eine multiplikative Basis für die Gruppe der Multiplikatoren von Substitutionen aus $\mathbf{A}(s)$, und S_1, S_2, \dots, S_t seien die zugehörigen Substitutionen. Setzen wir voraus, daß stets

$$-E = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \in G,$$

so hat man in den Substitutionen $-E, S_1, S_2, \dots, S_t$ eine Basis für $\mathbf{A}(s)$. Die Unabhängigkeit der Multiplikatoren $\lambda_1^2, \lambda_2^2, \dots, \lambda_t^2$ hat wegen $\lambda_\mu^{(\nu)} \lambda_\mu^{(\nu)} = 1$ für $\mu \leq t, \nu > t$ zur Folge, daß die Determinante der Matrix $(\log \lambda_\mu^{(\nu)2})$, $\mu, \nu \leq t$, nicht verschwindet:

$$(10) \quad |\log \lambda_\mu^{(\nu)2}| \neq 0 \quad (\mu, \nu \leq t).$$

Im folgenden benötigen wir eine feststehende Definition der analytischen

Funktion $(a_1\tau + a_2)^r$ für ein beliebiges Paar komplexer Zahlen $a_1, a_2 \neq 0, 0$, eine komplexe Veränderliche τ und $a_1\tau + a_2 \neq 0$. Wir setzen fest:

$$(11) \quad \begin{aligned} (a_1\tau + a_2)^r &= e^{r \log(a_1\tau + a_2)}, \\ \log(a_1\tau + a_2) &= \log|a_1\tau + a_2| + i \arg^*(a_1\tau + a_2), \\ \arg^*(a_1\tau + a_2) &= \begin{cases} \arg\left(\tau + \frac{a_2}{a_1}\right) - \arg \bar{a}_1 & \text{für } a_1 \neq 0, \\ \arg a_2 & \text{für } a_1 = 0, \end{cases} \\ -\pi < \arg \omega \leq \pi & \text{ für eine beliebige komplexe Zahl } \omega \neq 0. \end{aligned}$$

Man beachte, daß der Wert der Funktion $(a_1\tau + a_2)^r$ nicht nur von $a_1\tau + a_2$, sondern auch noch von a_1 abhängt. Ist $a_1 \neq 0$ und $\Im m \frac{a_2}{a_1} \geq 0$, so ist $(a_1\tau + a_2)^r$ im ganzen Bereich $\Im m \tau > 0$ regulär analytisch. Für ein reelles Zahlenpaar a_1, a_2 stimmt die Definition (11) mit der Regelung in ¹⁾ [Formel (7)] überein. In den weiteren Betrachtungen soll stets $\Im m \tau > 0$ und im Falle $a_1 \neq 0$ auch $\Im m \frac{a_2}{a_1} \geq 0$ vorausgesetzt werden. Sei S eine unimodulare reelle Matrix mit der zweiten Zeile $\underline{S} = (\gamma, \delta)$ und $(a_1, a_2)S = (a_1^*, a_2^*)$. Eine komplexe Matrix A werde so bestimmt, daß $\underline{A} = (a_1, a_2)$. Dann besteht die Transformationsformel

$$(12) \quad (a_1S\tau + a_2)^r = \sigma_0^{(r)}(A, S) \frac{(a_1^*\tau + a_2^*)^r}{(\gamma\tau + \delta)^r} \quad (\text{für } \Im m \tau > 0)$$

mit

$$(13) \quad \begin{aligned} \sigma_0^{(r)}(A, S) &= e^{2\pi i r w(A, S)}, \\ 2\pi w(A, S) &= \arg^*(a_1S\tau + a_2) - \arg^*(a_1^*\tau + a_2^*) + \arg^*(\gamma\tau + \delta). \end{aligned}$$

Für reelle Argumente hat H. Petersson die Funktion $w(A, S)$ bestimmt⁴⁾.

Nun zurück zu unserm Fixpunktsystem s ! Wir bestimmen eine Transformation

$$A = \begin{pmatrix} a_1 & a_2 \\ a_1' & a_2' \end{pmatrix} \quad \text{mit } |A^{(v)}| > 0 \text{ für } v \leq t, \quad -i|A^{(v)}| > 0 \text{ für } v > t,$$

welche \mathfrak{H} in den Bereich $\Im m \tau^{(v)} > 0$ für $v \leq t$ und $|\tau^{(v)}| < 1$ für $v > t$ überführt und dabei s in die spezielle Lage

$$(14) \quad A\tau_0 = 0, \quad A\tau_0' = \infty$$

transformiert. Die Determinantenbedingungen für A sind gleichwertig damit, daß $A^{(v)}$ für $v \leq t$ reell ist, für $v > t$ dagegen konjugiert komplexe Zeilen hat. In Anlehnung an ⁵⁾ untersuchen wir nun die Wirkung der Transformationen von $A(s)$ bei Anwendung auf die Hilfsfunktion

$$h(\tau) = h_A(\tau) = \prod_{v=1}^t (a_1^{(v)}\tau^{(v)} + a_2^{(v)})^{\frac{r}{2}} (a_1^{(v)}\tau^{(v)} + a_2^{(v)})^{\frac{r}{2}} \prod_{v=t+1}^n (a_1^{(v)}\tau^{(v)} + a_2^{(v)})^r$$

⁴⁾ H. Petersson, Zur analytischen Theorie der Grenzkreisgruppen I. Math. Annalen 115 (1938), S. 23–67, insbesondere S. 44, Satz 4.

⁵⁾ H. Petersson, Zur analytischen Theorie der Grenzkreisgruppen V. Math. Zeitschr. 44 (1938), S. 127–155.

mit dem Ziel, $h(\tau)$ als automorphe Form der Dimension $+r$ zu $\mathbf{A}(s)$ und einem gewissen Multiplikatorsystem v_0 zu erkennen. Es sei

$$D_\mu = \begin{pmatrix} \mu & 0 \\ 0 & \mu^{-1} \end{pmatrix} \text{ (allgemein), } T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad A' = -TA = \begin{pmatrix} -a'_1 & -a'_2 \\ a_1 & a_2 \end{pmatrix},$$

$$S = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathbf{A}(s), \quad S = A^{-1}D_\lambda A = A'^{-1}D_{\lambda'} A' \quad (\lambda\lambda' = 1),$$

$$g^{(\nu)}(\tau^{(\nu)}) = \begin{cases} (a_1^{(\nu)\prime} \tau^{(\nu)} + a_2^{(\nu)\prime})^{\frac{r}{2}} (a_1^{(\nu)} \tau^{(\nu)} + a_2^{(\nu)})^{\frac{r}{2}} & \text{für } \nu \leq t, \\ (a_1^{(\nu)\prime} \tau^{(\nu)} + a_2^{(\nu)\prime})^r & \text{für } \nu > t. \end{cases}$$

Vernachlässigen wir vorübergehend den Konjugiertenindex ν , so ergibt sich

1. für $\nu \leq t$:

$$g(S\tau) = (a'_1 A^{-1} D_\lambda A \tau + a'_2)^{\frac{r}{2}} (a_1 A'^{-1} D_{\lambda'} A' \tau + a_2)^{\frac{r}{2}}$$

$$= \sigma_0^{\left(\frac{r}{2}\right)}(A, S) \sigma_0^{\left(\frac{r}{2}\right)}(A', S) (\gamma\tau + \delta)^{-r} (\lambda' a'_1 \tau + \lambda' a'_2)^{\frac{r}{2}} (\lambda a_1 \tau + \lambda a_2)^{\frac{r}{2}}$$

$$= e^{\pi i r \operatorname{sgn}_+ \gamma \frac{1 - \operatorname{sgn} \lambda}{2}} (\gamma\tau + \delta)^{-r} g(\tau),$$

2. für $\nu > t$:

$$g(S\tau) = e^{r(\log|a'_1| - i \arg a'_1)} (S\tau - \tau'_0)^r = e^{r(\log|a'_1| - i \arg a'_1)} (S\tau - S\tau'_0)^r$$

$$= e^{r(\log|a'_1| - i \arg a'_1)} (\gamma\tau'_0 + \delta)^{-r} (\gamma\tau + \delta)^{-r} (\tau - \tau'_0)^r$$

$$= \lambda^{-r} (\gamma\tau + \delta)^{-r} g(\tau),$$

woraus erhellt, daß

$$(15) \quad h(S\tau) = v_0(S) N(\gamma\tau + \delta)^{-r} h(\tau) \quad \text{für } S \in \mathbf{A}(s)$$

mit

$$(16) \quad v_0(S) = \prod_{\nu=1}^t e^{\pi i r \operatorname{sgn}_+ \gamma^{(\nu)} \frac{1 - \operatorname{sgn} \lambda^{(\nu)}}{2}} \prod_{\nu=t+1}^n \lambda^{(\nu)-r}.$$

Für eine automorphe Form $f(\tau)$ vom Typus $\{\mathbf{G}, -r, v\}$ ist daher $g(\tau) = h(\tau)f(\tau)$ eine Form von der Dimension 0 zur Gruppe $\mathbf{A}(s)$ und zum Multiplikatorsystem v_0 . Dieses ist also ein abelscher Charakter für $\mathbf{A}(s)$. Die Invarianzeigenschaft von $g(\tau)$:

$$(17) \quad g(S\tau) = v(S) v_0(S) g(\tau) \quad \text{für } S \in \mathbf{A}(s)$$

wird uns die eingangs beschriebene Entwicklung für $f(\tau)$ liefern. Die „Orts-uniformisierenden“ werden schrittweise wie folgt eingeführt. Wir setzen $\tau_1 = A\tau$ und bestimmen aus dem Gleichungssystem

$$(18) \quad \log \tau_1^{(\nu)} = \sum_{\mu=1}^t \log \lambda_\mu^{(\nu)2} z_\mu \quad (\nu = 1, 2, \dots, t)$$

⁶⁾ Es ist $\operatorname{sgn}_+ \gamma = \operatorname{sgn} \gamma$ für $\gamma \neq 0$, $\operatorname{sgn}_+ 0 = 1$ gesetzt.

durch Umkehrung die Variablen

$$(19) \quad z_\mu = \sum_{\nu=1}^t m_\mu^{(\nu)} \log \tau_1^{(\nu)}.$$

Nach (10) ist das möglich. Ferner sei

$$(20) \quad \lambda_\mu^{(\nu)2} = e^{2\pi i q_\mu^{(\nu)}} \quad \text{für } \mu = 1, 2, \dots, t, \nu = t+1, t+2, \dots, n,$$

$$m_\mu^{(\nu)} = - \sum_{\varrho=1}^t m_\varrho^{(\nu)} q_\varrho^{(\mu)} \quad \text{für } \mu = t+1, t+2, \dots, n, \nu = 1, 2, \dots, t$$

und schließlich

$$(21) \quad \omega_\mu = e^{2\pi i z_\mu} = \prod_{\nu=1}^t (\tau_1^{(\nu)})^{2\pi i m_\mu^{(\nu)}} \quad \text{für } \mu \leq t,$$

$$\omega_\mu = \tau_1^{(u)} \prod_{\nu=1}^t e^{-2\pi i q_\nu^{(u)} z_\nu} = \tau_1^{(u)} \prod_{\nu=1}^t (\tau_1^{(\nu)})^{2\pi i m_\mu^{(\nu)}} \quad \text{für } \mu > t.$$

Der Bereich \mathfrak{H} wird beschrieben durch die Ungleichungen

$$0 < \Im \log \tau_1^{(\nu)} < \pi \quad \text{für } \nu \leq t \quad \text{und} \quad |\tau_1^{(\nu)}| < 1 \quad \text{für } \nu > t$$

oder in den „Ortsuniformierenden“ durch

$$(22) \quad e^{-2\pi^2} < \prod_{\mu=1}^t |\omega_\mu|^{\log \lambda_\mu^{(\nu)2}} < 1 \quad \text{für } \nu \leq t,$$

$$|\omega_\mu| \prod_{\nu=1}^t |\omega_\nu|^{q_\nu^{(u)}} < 1 \quad \text{für } \mu > t.$$

Das ist ein sogenannter Reinhardt'scher Körper mit dem Mittelpunkt $(0, 0, \dots, 0)$. Zu den Erzeugenden S_μ der Gruppe bestimme man nunmehr Zahlen κ_μ im Intervall $0 \leq \kappa_\mu < 1$ durch die Gleichungen

$$(23) \quad v(S_\mu) v_0(S_\mu) = e^{2\pi i \kappa_\mu} \quad \text{für } \mu = 1, 2, \dots, t.$$

Die Funktion

$$(24) \quad F(\tau) = e^{-\sum_{\mu=1}^t 2\pi i \kappa_\mu z_\mu} \quad g(\tau) = \prod_{\mu=1}^t \omega_\mu^{-\kappa_\mu} \cdot g(\tau)$$

erweist sich dann gegenüber den Substitutionen von \mathbf{A} (s) als absolut invariant:

$$(25) \quad F(S\tau) = F(\tau) \quad \text{für } S \subset \mathbf{A} \text{ (s)}.$$

Das ergibt sich sofort aus (17), wenn man beachtet, daß entsprechend einer Ersetzung von τ durch $S_\mu \tau$ die Variable z_μ in $z_\mu + 1$ übergeht, z_ϱ für $\varrho \neq \mu$ dagegen unverändert bleibt. Die Invarianz (25) besagt nichts anderes, als daß $F(\tau)$ eine im Bereich (22) eindeutige Funktion der ω_μ ist. Um das einzusehen, brauchen wir uns nur zu überlegen, wieweit ein Punkt τ durch die Variablen $\omega_1, \omega_2, \dots, \omega_n$ bestimmt ist. Da die $z_\mu \bmod 1$ bestimmt sind, genügt es offenbar, die Wirkung festzustellen, die durch die Abänderung etwa

von z_1 in $z_1 + 1$ hervorgerufen wird, während alle anderen z_μ festbleiben. Zunächst folgt nach (18), daß $\tau_1^{(v)}$ für $v \leq t$ den Faktor $\lambda_1^{(v)^2}$ aufnimmt; nach (20) und (21) gilt das auch für $v > t$. τ_1 durch $\lambda_1^2 \tau_1$ ersetzen heißt aber τ in $S_1 \tau$ überführen. Damit ist die Behauptung bewiesen. $F(\tau)$ ist also eine im Bereich (22) eindeutige und reguläre Funktion der ω_μ , als solche in eine Laurent-Reihe der Art

$$(26) \quad F(\tau) = \sum b_{k_1 k_2 \dots k_n} \omega_1^{k_1} \omega_2^{k_2} \dots \omega_n^{k_n}$$

mit den Summationsbedingungen

$$(27) \quad -\infty < k_\mu < \infty \quad \text{für } \mu \leq t, \quad 0 \leq k_\mu < \infty \quad \text{für } \mu > t$$

zu entwickeln. Für $f(\tau)$ ergibt sich damit die Entwicklung

$$(28) \quad h(\tau)f(\tau) = \sum b_{k_1 k_2 \dots k_n} \omega_1^{k_1 + \kappa_1} \dots \omega_t^{k_t + \kappa_t} \omega_{t+1}^{k_{t+1}} \dots \omega_n^{k_n}.$$

Das Ergebnis ist im Fall $t = 0$ trivial. Es sei daran erinnert, daß im Fall $t = 0$ das System s aus einem einzigen inneren Punkt τ_0 von \mathfrak{H} bestehen soll. $\mathcal{A}(s)$ bestehe dann aus den beiden Substitutionen $\pm E$. Die Entwicklungskoeffizienten $b_{k_1 k_2 \dots k_n}$ sind bei festem Punktsystem s noch abhängig von der Auswahl der Basis S_1, S_2, \dots, S_t , sowie der Transformation A . Die Art der Abhängigkeit läßt sich in beiden Fällen übersehen. Setzt man nämlich noch $\kappa_\mu = 0$ für $\mu > t$ und

$$(29) \quad \sum_{\mu=1}^n m_\mu^{(v)} (k_\mu + \kappa_\mu) = j^{(v)} \quad \text{für } v \leq t,$$

so wird

$$(30) \quad \omega_1^{k_1 + \kappa_1} \dots \omega_t^{k_t + \kappa_t} \omega_{t+1}^{k_{t+1}} \dots \omega_n^{k_n} = \prod_{v=1}^t (\tau_1^{(v)})^{2\pi i j^{(v)}} \prod_{v=t+1}^n (\tau_1^{(v)})^{k_v}.$$

Die umbenannten Koeffizienten

$$(31) \quad c(j^{(1)}, \dots, j^{(t)}, k_{t+1}, \dots, k_n) = b_{k_1 k_2 \dots k_n}$$

sind dann offenbar von der Basis S_μ nicht abhängig. Wählt man andererseits eine andere Transformation B an Stelle von A , die das gleiche leistet wie A , so ist notwendig

$$B = \varrho D_\lambda A,$$

wobei $\varrho > 0$, $\lambda_0^{(v)}$ für $v \leq t$ reell und für $v > t$ vom Betrag 1 ist. Bei der Ersetzung von A durch B nimmt $h(\tau)$ einen von ϱ , λ_0 und A abhängigen Faktor auf, τ_1 multipliziert sich mit λ_0^2 , so daß sich schließlich nach (30) auch $b_{k_1 k_2 \dots k_n}$ nur um einen von A , ϱ , λ_0 und (k) abhängigen Faktor ändert. In diesem Sinne ist die Entwicklung (28) von $f(\tau)$ zum Punktsystem s eindeutig. Schließlich soll noch bewiesen werden, daß die Entwicklung (28) gegenüber Transformation von $f(\tau)$ mit einer beliebigen reellen unimodularen Substitution S invariant

ist. Trägt man $S\tau$ an Stelle von τ in (28) ein, so gehen die „Ortsuniformisierenden“ ω_μ in ein System von „Ortsuniformisierenden“ zum Punktsystem $S^{-1}(s)$ über. Die linke Seite von (28): $h(\tau)f(\tau) = h_A(\tau)f(\tau)$ verwandelt sich in

$$(32) \quad h_A(S\tau)f(S\tau) = \zeta^{(r)}(A, S) h_{AS}(\tau) N(\gamma\tau + \delta)^{-r} f(S\tau) \\ = \zeta^{(r)}(A, S) h_{AS}(\tau) f^S(\tau) \text{ [Bezeichnung s. 1)]}$$

mit

$$(33) \quad \zeta^{(r)}(A, S) = \prod_{v=1}^t \sigma_0^{(r)} \left(A^{(v)}, S^{(v)} \right) \sigma_0^{(r)} \left(A^{(v)}, S^{(v)} \right) \prod_{v=t+1}^n \sigma_0^{(r)} \left(A^{(v)}, S^{(v)} \right).$$

Die Entwicklungskoeffizienten der Form $f^S(\tau) \in \{S^{-1}GS, -r, v^S\}$ zum Punktsystem $S^{-1}(s)$ stimmen also mit denen der Form $f(\tau)$ zum Punktsystem s im wesentlichen überein.

§ 2.

Poincarésche Reihen. Transformationsformeln und Konvergenz.

Zu jeder Substitution $S \in G$ gibt es auf Grund von (28) bis (32) und

$$f^S(\tau) = v(S)f(\tau)$$

eine Entwicklung

$$(34) \quad f(\tau) = \sum_{k_1, k_2, \dots, k_n} b_{k_1, k_2, \dots, k_n} \frac{\prod_{v=1}^t (A^{(v)} S^{(v)} \tau^{(v)})^{2\pi i f^{(v)}} \prod_{v=t+1}^n (A^{(v)} S^{(v)} \tau^{(v)})^{k_v}}{\zeta(A, S) v(S) h_{AS}(\tau)}.$$

Wir zerlegen nun G in Linksrestklassen nach $A(s)$:

$$G = \sum_{S^*} A(s) S^*$$

und vereinigen alle Transformationen AS^* zu einer Menge $\mathfrak{S}(A, G)$. Ausgenommen, daß $t=0$ und τ_0 elliptischer Fixpunkt von G ist, gelangt man zu einer solchen Menge $\mathfrak{S}(A, G)$ auch, wenn man aus der Menge AG ein vollständiges System von Substitutionen mit paarweise nicht assoziierten zweiten Zeilen aussondert. Dabei heißen die zweiten Zeilen zweier Substitutionen S_1 und S_2 assoziiert, wenn es ein Zahlensystem $\lambda = \{\lambda^{(v)}\}$ gibt, so daß $\underline{S_2} = \lambda \underline{S_1}$.

Im Fall $t=0$ bedeutet die Auswahl von $\mathfrak{S}(A, G)$, daß für jedes $S \in G$ entweder AS oder $-AS$ in AG ausgezeichnet wird. Über (34) gelangen wir jetzt zu einer Poincaréschen Reihe, indem wir in (34) formal $b_{q_1, q_2, \dots, q_n} = 1$, alle anderen Koeffizienten gleich 0 setzen und über $AS \in \mathfrak{S}(A, G)$ summieren:

$$(35) \quad \Xi_{-r}(\tau) = \Xi_{-r}(\tau, v, A, G, (q)) \\ = \sum_{AS \in \mathfrak{S}(A, G)} \frac{\prod_{v=1}^t (A^{(v)} S^{(v)} \tau^{(v)})^{2\pi i p^{(v)}} \prod_{v=t+1}^n (A^{(v)} S^{(v)} \tau^{(v)})^{q_v}}{\zeta(A, S) v(S) h_{AS}(\tau)}.$$

Die Zahlen $p^{(r)}$ haben für das System (q) dieselbe Bedeutung wie $j^{(r)}$ für (k) . Zunächst wollen wir uns davon überzeugen, daß $\Xi_{-r}(\tau)$ von der Auswahl des Systems $\mathfrak{S}(A, G)$ nicht abhängt. Wir ersetzen im allgemeinen Summenglied S durch $S_0 S$, wobei $S_0 = A^{-1} D_A A \subset A(s)$ und etwa

$$\lambda^2 = \lambda_1^{2a_1} \lambda_2^{2a_2} \dots \lambda_r^{2a_r}.$$

Da nun $AS\tau$ in $\lambda^2 AS\tau$ übergeht, so nimmt der Zähler in (35) folgenden Faktor auf:

$$\begin{aligned} & e^{\frac{2\pi i}{t} \sum_{\mu \leq t} \log \lambda^{(r)2} p^{(r)}} \prod_{r > p} \lambda^{(r)2 q_r} \\ &= e^{\frac{2\pi i}{t} \sum_{\mu \leq t} (q_\mu + \kappa_\mu) a_\mu - \sum_{\mu > t} q_\mu \sum_{\mu \leq t} \frac{2\pi i q^{(\mu)} a}{q}} \prod_{r > p} \lambda^{(r)2 q_r} \\ &= e^{\frac{2\pi i}{t} \sum_{\mu \leq t} \kappa_\mu a_\mu} = v(S_0) v_0(S_0). \end{aligned}$$

Der Nenner

$$\zeta(A, S) v(S) h_{AS}(\tau) = v(S) N(\gamma\tau + \delta)^r h_A(S\tau) = v(S) h_A^S(\tau)$$

$[h_A(\tau)$ ist eine Form der Dimension $+r$] dagegen wird ersetzt durch

$$\begin{aligned} v(S_0 S) h_A^{S_0 S}(\tau) &= v(S_0 S) \sigma^{(-r)}(S_0, S) (h_A^{S_0}(\tau))^S \\ &= v(S) h_A^S(\tau) \cdot v(S_0) v_0(S_0) \quad [\text{vgl. } ^1), \text{ Formeln (19), (20)}] \end{aligned}$$

und nimmt also ebenfalls den Faktor $v(S_0) v_0(S_0)$ auf, q. e. d. Zur Ableitung der Transformationsformeln für Ξ_{-r} benötigen wir noch folgende Beziehung für $\zeta^{(r)}(A, S)$. Für zwei reelle unimodulare Matrizen S, L folgt zunächst aus dem Assoziativgesetz für die Matrizenmultiplikation, angewendet auf (12), daß

$$\sigma_0^{(r)}(A, SL) \sigma_0^{(r)}(S, L) = \sigma_0^{(r)}(AS, L) \sigma_0^{(r)}(A, S)$$

oder

$$(36) \quad w(A, SL) + w(S, L) = w(AS, L) + w(A, S).$$

Zu der bekannten Formel

$$(37) \quad \sigma^{(r)}(A, SL) \sigma^{(r)}(S, L) = \sigma^{(r)}(AS, L) \sigma^{(r)}(A, S)$$

gesellt sich dann auf Grund von (36) die Relation

$$(38) \quad \zeta^{(r)}(A, SL) \sigma^{(r)}(S, L) = \zeta^{(r)}(AS, L) \zeta^{(r)}(A, S).$$

Wir transformieren nun $\Xi_{-r}(\tau)$ mit einer beliebigen reellen unimodularen Substitution L . $\underline{L} = (\gamma, \delta)$ sei die zweite Zeile von L . Mit $B = AS$ wird dann

$$(39) \quad \Xi_{-r}^L(\tau) = \sum_{B \in \mathfrak{S}(A, G)} \frac{\prod_{r=1}^t (B^{(r)} L^{(r)} \tau^{(r)})^{2\pi i p^{(r)}} \prod_{r=t+1}^n (B^{(r)} L^{(r)} \tau^{(r)})^{q_r}}{\zeta(A, S) v(S) N(\gamma\tau + \delta)^r h_B(L\tau)}.$$

Durch eine einfache Umformung geht der Nenner des allgemeinen Summanden in (39) zunächst in

$$(40) \quad \zeta(A, S) \zeta(B, L) v(S) h_{BL}(\tau)$$

über. Wir unterscheiden jetzt zwei Fälle:

1. Es sei $L \subset G$. Mit Hilfe von (38) verwandelt man den Nenner (40) in

$$\zeta(A, SL) v(SL) h_{BL}(\tau) \overline{v(L)}.$$

Da die Matrizenysteme $\mathfrak{S}(A, G)$ und $\mathfrak{S}(A, G)L$ im Hinblick auf die Ξ -Reihen gleichwertig sind, so erhält man aus (39) die Transformationsformel

$$(41) \quad \Xi_{-,r}^L(\tau, v, A, G, (q)) = v(L) \Xi_{-,r}(\tau, v, A, G, (q)) \quad \text{für } L \subset G.$$

2. Es sei L beliebig. Mit dem Nenner (40) wird dann folgende Umformung vorgenommen:

$$S^* = L^{-1}SL, \quad BL = ALS^*, \quad \sigma = \sigma^{(r)}, \quad \zeta = \zeta^{(r)},$$

$$\zeta(A, S) \zeta(B, L) v(S) h_{BL}(\tau)$$

$$\begin{aligned} &= \frac{\zeta(A, S)}{\zeta(AL, L^{-1}S)} \zeta(AL, S^*) \sigma(L^{-1}S, L) v(S) h_{BL}(\tau) \\ &= \zeta(A, L) \zeta(AL, S^*) \frac{\sigma(L^{-1}S, L)}{\sigma(L, L^{-1}S)} v(S) h_{BL}(\tau) \\ &= \zeta(A, L) \zeta(AL, S^*) v(S) \frac{\sigma(S, L)}{\sigma(L, S^*)} h_{BL}(\tau) \\ &= \zeta(AL, S^*) v^L(S^*) h_{ALS^*}(\tau) \cdot \zeta(A, L). \end{aligned}$$

Durchläuft B alle Transformationen von $\mathfrak{S}(A, G)$, so durchläuft $B^* = BL$ ein System vom Typus $\mathfrak{S}(AL, L^{-1}GL)$. (39) liefert dann

$$\Xi_{-,r}^L(\tau) = \frac{1}{\zeta(A, L)} \sum_{B^* \in \mathfrak{S}(AL, L^{-1}GL)} \frac{\prod_{v=1}^t (B^{*(v)} \tau^{(v)})^{2\pi i p^{(v)}} \prod_{v=t+1}^n (B^{*(v)} \tau^{(v)})^{q_v}}{\zeta(AL, S^*) v^L(S^*) h_{B^*}(\tau)}.$$

Die unendliche Reihe auf der rechten Seite dieser Gleichung ist offenbar eine Poincarésche Reihe vom Ξ -Typus mit den Argumenten $v^L, AL, L^{-1}GL, (q)$. Für eine beliebige reelle unimodulare Substitution L besteht also die Transformationsgleichung

$$(42) \quad \Xi_{-,r}^L(\tau, v, A, G, (q)) = \frac{1}{\zeta^{(r)}(A, L)} \Xi_{-,r}(\tau, v^L, AL, L^{-1}GL, (q)).$$

Speziell für $L = \lambda A^{-1}A_t$ mit $\lambda^2 = |A_t^{-1}A|$, $\lambda > 0$ und der festen Transformation A_t , definiert durch

$$(43) \quad A_t^{(v)} = E^{(v)} \quad \text{für } v \leq t, \quad A_t^{(v)} = \begin{pmatrix} 1 & -i \\ & i \end{pmatrix} \quad \text{für } v > t,$$

erhalten wir rechtsseitig von (42) eine Poincarésche Reihe $\Xi_{-,r}(\tau, v^L, \lambda A_t, L^{-1}GL, (q))$ zu dem speziellen Punktsystem

$$(44) \quad \tau_0^{(v)} = 0, \tau_0^{(v)'} = \infty \quad \text{für } v \leq t, \quad \tau_0^{(v)} = i, \tau_0^{(v)'} = -i \quad \text{für } v > t.$$

In dieser Reihe kann λA_t durch A_t ersetzt werden, wenn man $N \lambda^{-r}$ als Faktor anbringt, wie sich aus

$$\begin{aligned} \zeta(\lambda A_t, S) h_{\lambda A_t S}(\tau) &= N(\gamma\tau + \delta)^r h_{\lambda A_t}(S\tau), \quad \underline{S} = (\gamma, \delta), \\ (45) \quad h_{\lambda A_t}(S\tau) &= \prod_{v=1}^t \lambda^{(v)\frac{r}{2}} (\lambda^{(v)} S^{(v)} \tau^{(v)})^{\frac{r}{2}} \prod_{v=t+1}^n (\lambda^{(v)} S^{(v)} + \lambda^{(v)} i)^r = N \lambda^r h_{A_t}(S\tau) \end{aligned}$$

ergibt. Als spezielle Folge von (42) notieren wir

$$(46) \quad \Xi_{-,r}^L(\tau, v, A, G, (q)) = \frac{1}{\zeta^{(v)}(A, L) N \lambda^r} \Xi_{-,r}(\tau, v^L, A_t, L^{-1}GL, (q))$$

mit $L = \lambda A^{-1}A_t, \quad \lambda^2 |A^{-1}A_t| = 1, \quad \lambda > 0.$

Die Gruppe $L^{-1}GL$ kann natürlich nicht in der Hilbertschen Modulgruppe liegen, wenn sie das hyperbolische Fixpunktsystem (44) besitzt. Wenn in den folgenden Konvergenzbeweisen für die Poincaréschen Reihen dennoch von solchen Gruppen gesprochen wird, so kann das nicht stören; denn sie entstehen immer durch Transformation aus Untergruppen G der Hilbertschen Modulgruppe und alle Eigenschaften von G , auf welchen unsere Überlegungen basieren, sind gegenüber Transformationen von \mathfrak{H} in sich invariant. Hat das Fixpunktsystem s für eine Gruppe G die spezielle Lage (44) und wählen wir $A = A_t$, so nehmen die Poincaréschen Reihen, wie aus (45) für $\lambda = 1$ hervorgeht, die einfache Gestalt

$$(47) \quad \Xi_{-,r}(\tau, v, A_t, G, (q)) = \sum_{s \in \mathfrak{S}_s(G)} \frac{\prod_{v=1}^t (S^{(v)} \tau^{(v)})^{2\pi i p^{(v)} - \frac{r}{2}} \prod_{v=t+1}^n \left(\frac{S^{(v)} \tau^{(v)} - i}{S^{(v)} \tau^{(v)} + i} \right)^{q_v}}{v(S) N(\gamma\tau + \delta)^r \prod_{v=t+1}^n (S^{(v)} \tau^{(v)} + i)^r}$$

an, wobei $\mathfrak{S}_s(G)$ ein volles System von Repräsentanten der Linksrestklassen von G nach $A(s)$ darstellt.

Wir beweisen nun die absolute Konvergenz der Poincaréschen Reihen $\Xi_{-,r}(\tau)$ für $r > 2$, $|v| = 1$, ein beliebiges ganzzahliges Exponentensystem (q) mit $q_\mu \geq 0$ für $\mu > t$ und alle Punkte von \mathfrak{H} . Da der Zähler des allgemeinen Gliedes der Reihe (35) bei festem Exponentensystem (q) nach (22), (30) beschränkt ist, also dem Betrage nach etwa kleiner als C_0 ist, so besitzt $\frac{1}{C_0} \Xi_{-,r}(\tau)$ die Majorante

$$(48) \quad \Omega(\tau, r, A, G) = \sum_{B \in \mathfrak{S}(A, G)} \frac{1}{|h_B(\tau)|}.$$

Der Beweis für die Konvergenz von Ω wird entsprechend den drei Fällen $t = n$, $t = 0$, $0 < t < n$ verschieden geführt, im ersten und wichtigsten Fall mit Hilfe der nichteuklidischen Maßbestimmung in §3). Durch einen eleganten Ansatz, der auf Herrn Petersson zurückgeht, wird dabei in der Beweisführung eine erhebliche Vereinfachung erreicht im Vergleich etwa zu den analogen Bemühungen¹⁾ für Poincarésche Reihen zu den parabolischen Spitzen.

1. $t = n$. Es ist $|h_B(\tau)| = |h_{B_A-1}(A\tau)| N|a'_1\tau + a'_2|^r$, also

$$\Omega(\tau, r, A, G) = N|a'_1\tau + a'_2|^{-r} \Omega(A\tau, r, E, AGA^{-1}).$$

Ohne Einschränkung der Allgemeinheit können wir daher in (48) $A = A_n = E$, d. h. s in der speziellen Lage (44) voraussetzen. Die Erzeugenden S_μ von $A(s)$ haben dann die Gestalt

$$S_\mu = \begin{pmatrix} \lambda_\mu & 0 \\ 0 & \lambda_\mu^{-1} \end{pmatrix} \quad \text{für } \mu = 1, 2, \dots, n.$$

Man bestimme x_1, x_2, \dots, x_n aus dem Gleichungssystem

$$(49) \quad \log |\tau^{(v)}| = x_1 \log \lambda_1^{(v)2} + x_2 \log \lambda_2^{(v)2} + \dots + x_n \log \lambda_n^{(v)2} \quad (v = 1, 2, \dots, n).$$

Im Raum der Punkte mit den affinen Koordinaten (x_1, x_2, \dots, x_n) ist $A(s)$ eine Translationsgruppe mit dem Fundamentalbereich

$$(50) \quad -\frac{1}{2} < x_j \leq \frac{1}{2} \quad \text{für } j = 1, 2, \dots, n.$$

τ heißt reduziert, wenn der Bildpunkt (x_1, x_2, \dots, x_n) im Fundamentalbereich (50) liegt. Zu einem festen Punkt τ^* ist dann ein System $\mathfrak{S}_0 = \mathfrak{S}(E, G)$ durch die Vorschrift eindeutig bestimmt, daß $L\tau^*$ für alle $L \in \mathfrak{S}_0$ reduziert sein soll. Nach (50) kann eine positive Konstante c_1 bestimmt werden, so daß

$$(51) \quad \frac{1}{c_1} < |L\tau^*| < c_1 \quad \text{für } L \in \mathfrak{S}_0.$$

\mathfrak{R}_0 sei eine nichteuklidische Kugel um einen vorgegebenen Punkt τ^* mit dem Radius $\varrho_0 > 0$, dieser so klein, daß für $L \in G$, $L\tau^* \neq \tau^*$ die Punktmengen \mathfrak{R}_0 , $L\mathfrak{R}_0$ einen leeren Durchschnitt haben. Die folgenden Abschätzungen entnimmt man leicht aus ⁵⁾ S. 144–145:

$$\tau^* = x^* + iy^*, \tau = x + iy \in \mathfrak{R}_0, L = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathfrak{S}_0, L\tau^* = x_L^* + iy_L^*,$$

$$L\tau = x_L + iy_L,$$

$$(52) \quad |L\tau^*|e^{-\varrho_0} \leq |L\tau| \leq |L\tau^*|e^{\varrho_0},$$

$$e^{-2\varrho_0 n} \leq N\left(\frac{y_L}{y} \frac{y^*}{y}\right) = N\left(\frac{y_L}{y} \frac{y^*}{y_L}\right) = \frac{N|\gamma\tau^* + \delta|^2}{N|\gamma\tau + \delta|^2} \leq e^{2\varrho_0 n}.$$

Für einen beliebigen Punkt $\tau \in \mathfrak{R}_0$, $c_2 = c_1 e^{\varrho_0}$ und $L \in \mathfrak{S}_0$ ist daher nach (51)

$$\frac{1}{c_2} < |L\tau| < c_2$$

und folglich

$$\Omega(\tau, r, E, G) < c_2^{\frac{nr}{2}} \sum_{L \in \mathfrak{G}_0} \frac{1}{N|\gamma\tau + \delta|^r}.$$

Die Konvergenz der rechts stehenden Summe ergibt sich nun leicht aus einer Integralabschätzung:

$$\begin{aligned} V(L\mathfrak{R}_0) &= \int \cdots \int_{\substack{\tau_1 = L\tau \\ \tau \in \mathfrak{R}_0}} N(y_1^{\frac{r}{2}-2} dx_1 dy_1) = \int \cdots \int_{\tau \in \mathfrak{R}_0} N|\gamma\tau + \delta|^{-r} N(y^{\frac{r}{2}-2} dx dy) \\ &\geq e^{-r\varrho_0 n} N|\gamma\tau^* + \delta|^{-r} V(\mathfrak{R}_0) \quad (\tau_1 = x_1 + iy_1), \\ \sum_{L \in \mathfrak{G}_0} N|\gamma\tau^* + \delta|^{-r} &\leq \frac{e^{r\varrho_0 n}}{V(\mathfrak{R}_0)} \sum_{L \in \mathfrak{G}_0} V(L\mathfrak{R}_0) \\ &\leq \frac{l^* e^{r\varrho_0 n}}{V(\mathfrak{R}_0)} \int \cdots \int_{\substack{|x| \leq c_2 \\ 0 \leq y \leq c_2}} N(y^{\frac{r}{2}-2} dx dy) = \frac{l^* e^{r\varrho_0 n} (2c_2)^n c_2^{\left(\frac{r}{2}-1\right)n}}{V(\mathfrak{R}_0) \left(\frac{r}{2}-1\right)^n}. \end{aligned}$$

l^* ist die Anzahl der Substitutionen aus G , welche τ^* festlassen. Im ersten Fall ist also die Konvergenz bewiesen.

2. $t = 0$. Wir bestimmen ein hyperbolisches Fixpunktsystem s^* für G mit 2^n reellen Punkten, die man durch Mischung der Konjugierten von τ_0^* , $\tau_0^{*\prime}$ erhalten möge, sowie eine reelle unimodulare Transformation A^* mit der Eigenschaft

$$A^* \tau_0^* = 0, \quad A^* \tau_0^{*\prime} = \infty.$$

Dann ist wieder

$$|h_B(\tau)| = |h_{BA^{*-1}}(A^*\tau)| N|a_1^* \tau + a_2^*|^r$$

und

$$\Omega(\tau, r, A, G) = N|a_1^* \tau + a_2^*|^{-r} \Omega(A^*\tau, r, AA^{*-1}, A^*GA^{*-1}).$$

Die Gruppe A^*GA^{*-1} hat das spezielle hyperbolische Fixpunktsystem s_0 : $\tau^{(v)} = 0$ oder ∞ für $v = 1, 2, \dots, n$. Wir können uns also beim Konvergenzbeweis auf diesen Fall beschränken und wollen annehmen, daß bereits G die genannte Eigenschaft hat. Offenbar ist $A(s) \subset A(s_0)$; denn $A(s)$ enthält im vorliegenden Fall nur die Substitutionen $\pm E$. Man bestimme die Zerlegungen in Linksrestklassen

$$G = \sum_{L^*} A(s_0)L^*, \quad A(s_0) = \sum_{S^*} A(s)S^*$$

und hat dann in der Gesamtheit aller Substitutionen L^* ein System $\mathfrak{S}(E, G)$, in der Gesamtheit aller möglichen Produkte $B = S^*L^*$ ein System $\mathfrak{S}(A, G)$. Die Gesamtheit der Substitution S^* möge mit der Menge der Potenzprodukte

$$S_1^{\alpha_1} S_2^{\alpha_2} \dots S_n^{\alpha_n}, \quad S_\mu = \begin{pmatrix} \lambda_\mu & 0 \\ 0 & \lambda_\mu^{-1} \end{pmatrix}, \quad \mu = 1, 2, \dots, n,$$

übereinstimmen. Die Transformationen S_μ sollen mit $-E$ die Gruppe $A(s_0)$ erzeugen. Nach diesen Vorbereitungen wird

$$(53) \quad \Omega(\tau, r, A, G) = \sum_{B \in \mathfrak{S}(A, G)} \frac{1}{|h_B(\tau)|} \\ = N|a'_1|^{-r} \sum_{S \in \mathfrak{S}(E, G)} \frac{1}{N|\gamma\tau + \delta|^r} \sum_i \frac{1}{N|\lambda S\tau - \lambda^{-1}\bar{\tau}_0|^r},$$

$$S = (\gamma, \delta),$$

wobei λ alle möglichen Potenzprodukte $\lambda_1^{a_1} \lambda_2^{a_2} \dots \lambda_n^{a_n}$ durchläuft. Um \sum_i gleichmäßig in S abschätzen zu können, denken wir uns zu einem vorgegebenen Punkt τ^* das System $\mathfrak{S}(E, G)$ derart bestimmt, daß $S\tau^*$ für alle $S \in \mathfrak{S}(E, G)$ bezüglich $A(s_0)$ reduziert ist. Nach (51) gilt dann $\frac{1}{c_1} < |S\tau^*| < c_1$ mit einer geeigneten positiven Konstanten c_1 , woraus sich ergibt, daß

$$(54) \quad |\lambda S\tau^* - \lambda^{-1}\bar{\tau}_0| \geq c_2|\lambda + \lambda^{-1}|$$

für alle $S \in \mathfrak{S}(E, G)$, beliebiges $\lambda \neq 0$, $\Im \tau_0 > 0$ und eine geeignete positive Konstante c_2 , die nur von c_1 und τ_0 abhängt. Setzt man nämlich $\lambda^* = \frac{\lambda}{\lambda + \lambda^{-1}}$, so wird behauptet, daß $|\lambda^* S\tau^* - (1 - \lambda^*)\bar{\tau}_0| \geq c_2 > 0$. Das ist in der Tat richtig; denn auf der abgeschlossenen Mannigfaltigkeit $\frac{1}{c_1} \leq |S\tau^*| \leq c_1$, $\Im S\tau^* \geq 0$, $0 \leq \lambda^* \leq 1$ hat $|\lambda^* S\tau^* - (1 - \lambda^*)\bar{\tau}_0|$ ein positives Minimum, weil $\lambda^* S\tau^* = (1 - \lambda^*)\bar{\tau}_0$ mit $\Im \tau_0 > 0$ nicht verträglich ist. Eine brauchbare Abschätzung für \sum_i hat man nun nach (54) in

$$\sum_i N|\lambda S\tau^* - \lambda^{-1}\bar{\tau}_0|^{-r} \leq c_2^{-r} n \sum_i \frac{1}{N|\lambda + \lambda^{-1}|^r},$$

Die Summe auf der rechten Seite dieser Ungleichung konvergiert bestimmt dann, wenn das n -fache Integral

$$\int_{-\infty}^{\infty} \dots \int \frac{dx_1 dx_2 \dots dx_n}{N(|\lambda_1|^{x_1} |\lambda_2|^{x_2} \dots |\lambda_n|^{x_n} + |\lambda_1|^{-x_1} |\lambda_2|^{-x_2} \dots |\lambda_n|^{-x_n})^r}$$

existiert. Das ist aber sofort zu erkennen, wenn man es auf die Veränderlichen

$$z^{(\nu)} = \sum_{\mu=1}^n x_\mu \log |\lambda_\mu^{(\nu)}|, \quad \nu = 1, 2, \dots, n,$$

umschreibt. Bezeichnet R den Betrag der Determinante $|\log |\lambda_\mu^{(\nu)}||$, so erhält man nämlich

$$\frac{1}{R} \int_{-\infty}^{\infty} \dots \int N \left(\frac{dz}{(e^z + e^{-z})^r} \right) = \frac{1}{R} \left(\int_{-\infty}^{\infty} \frac{dz}{(e^z + e^{-z})^r} \right)^n.$$

Zum Konvergenzbeweis für $\Omega(\tau^*, r, A, G)$ benötigen wir jetzt nach (53) nur noch die Konvergenz von $\sum_{s \in \mathfrak{S}(A, G)} N |\gamma \tau^* + \delta|^{-r}$. Der Beweis hierfür ist aber im ersten Fall ($t = n$) geführt. Damit ist auch der zweite Fall erledigt.

3. $0 < t < n$. Es sei $L = \mu A^{-1} A_t$, $\mu^2 |A^{-1} A_t| = 1$, $\mu > 0$. Analog wie im ersten Fall zeigen die Gleichungen

$$|h_B(\tau)| = |h_{BL}(L^{-1}\tau)| N |-c\tau + a|^r, \quad L^{-1} = (-c, a),$$

$$\Omega(\tau, r, A, G) = N |-c\tau + a|^{-r} N \mu^{-r} \Omega(L^{-1}\tau, r, A_t, L^{-1}GL),$$

daß wir uns darauf beschränken dürfen, das hyperbolische Fixpunktsystem s von G in der speziellen Lage (44) anzunehmen. Die ersten t Konjugierten der Erzeugenden S_μ von $A(s)$ sind dann von der Gestalt

$$\begin{pmatrix} \lambda_\mu^{(v)} & 0 \\ 0 & \lambda_\mu^{(v)-1} \end{pmatrix} \quad \text{für } v \leq t, \mu \leq t.$$

Ein Punkt τ von \mathfrak{S} wird reduziert bezüglich $A(s)$ genannt, wenn die aus dem Gleichungssystem

$$\log |\tau^{(v)}| = x_1 \log \lambda_1^{(v)2} + x_2 \log \lambda_2^{(v)2} + \dots + x_t \log \lambda_t^{(v)2} \\ (v = 1, 2, \dots, t)$$

bestimmten Zahlen x_1, x_2, \dots, x_t im Intervall

$$-\frac{1}{2} < x_\mu \leq \frac{1}{2} \quad \text{für } \mu = 1, 2, \dots, t$$

liegen. Es ist leicht zu sehen, daß jedes volle System nach $A(s)$ äquivalenter Punkte genau einen reduzierten enthält. Das System $\mathfrak{S}(A_t, G) = A_t \mathfrak{S}_s(G)$ wird jetzt nach Wahl eines Punktes τ^* wieder festgelegt durch die Forderung, daß $L\tau^*$ für alle $L \in \mathfrak{S}_s(G)$ reduziert sein soll. Diesmal ergibt sich nur

$$\frac{1}{c_1} < |L^{(v)} \tau^{*(v)}| < c_1 \quad \text{für } v = 1, 2, \dots, t$$

und eine geeignete positive Konstante c_1 . Für eine Substitution $L \in \mathfrak{S}_s(G)$ mit der zweiten Zeile (γ, δ) gelten daher die Ungleichungen

$$\begin{aligned} |h_{A_t L}(\tau^*)| &= N |\gamma \tau^* + \delta|^r |h_{A_t}(L\tau^*)| \\ &= N |\gamma \tau^* + \delta|^r \prod_{v=1}^t |L^{(v)} \tau^{*(v)}|^{\frac{r}{2}} \prod_{v=t+1}^n |L^{(v)} \tau^{*(v)} + i|^r \\ &> N |\gamma \tau^* + \delta|^r c_1^{-\frac{r}{2}} \prod_{v=t+1}^n |L^{(v)} \tau^{*(v)} + i|^r \\ &> c_1^{-\frac{r}{2}} (c_1 + 1)^{-r(n-t)} N |\gamma \tau^* + \delta|^r |L\tau^* + i|^r. \end{aligned}$$

Man erhält damit

$$\Omega(\tau^*, r, A_n, G) < c_1^{\frac{r}{2}} (c_1 + 1)^{r(n-1)} \sum_{L \in \mathfrak{S}_n(G)} \frac{1}{N |\gamma \tau^* + \delta|^r |L \tau^* + i|^r}.$$

Der Nenner des allgemeinen Gliedes ist offenbar gleich

$$|h_{A_0 L}(\tau^*)|.$$

Da ferner $A_0 \in \mathfrak{S}_n(G)$ zu einem System $\mathfrak{S}(A_0, G)$ ergänzt werden kann, so erhält man schließlich

$$\Omega(\tau^*, r, A_n, G) < c_1^{\frac{r}{2}} (c_1 + 1)^{r(n-1)} \Omega(\tau^*, r, A_0, G).$$

Die Reihe $\Omega(\tau^*, r, A_0, G)$ konvergiert, wie die vorangehende Betrachtung (zu $t = 0$) ergab. Damit ist alles bewiesen.

Die absolute Konvergenz der Poincaréschen Reihen ist erkannt. Es ist jedoch für alles Folgende von Bedeutung, daß die Ξ -Reihen unter den oben angegebenen Voraussetzungen in jedem Bereich der Art

$$(55) \quad |x^{(v)}| \leq c, \quad y^{(v)} \geq \varepsilon > 0 \quad \text{für } v = 1, 2, \dots, n$$

nicht nur absolut, sondern auch gleichmäßig konvergieren. Diese Tatsache beruht im wesentlichen auf Überlegungen, die zu den folgenden beiden Hilfsätzen führen.

Hilfssatz 1. Für ein Paar reeller Zahlen γ, δ und $\tau = x + iy$ mit $|x| \leq c$, $y \geq \varepsilon > 0$ sowie eine geeignete positive Konstante $C_1 = C_1(c, \varepsilon)$ ist

$$|\gamma \tau + \delta| \geq C_1 |\gamma i + \delta|.$$

[Beweis siehe ⁵⁾ S. 147, Hilfssatz 3].

Hilfssatz 2. Es sei $S = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ eine reelle Matrix mit positiver Determinante, τ_0 eine Zahl mit positivem Imaginärteil. Für $\tau = x + iy$ mit $|x| \leq c$, $y \geq \varepsilon > 0$ und eine geeignete positive Konstante $C_2 = C_2(c, \varepsilon, \tau_0)$ ist dann

$$|\gamma \tau + \delta| |S \tau + \tau_0| \geq C_2 |\gamma i + \delta| |S i + \tau_0|.$$

Zum Beweis bestimmen wir bei vorgegebenem τ eine Zahl $\lambda > 0$ so, daß $S = D_1 S_0$, $|S_0 \tau| = 1$. Es ist also $\lambda^2 = |S \tau|$. (γ_0, δ_0) sei die zweite Zeile von S_0 . Wir setzen $\tau_0 = |\tau_0| e^{i\varphi}$, $\lambda_1 = \frac{\lambda}{\lambda + \lambda^{-1} |\tau_0|}$ und erhalten auf Grund von Hilfssatz 1

$$\begin{aligned} |\gamma \tau + \delta| |S \tau + \tau_0| &= |\gamma_0 \tau + \delta_0| |\lambda S_0 \tau + \lambda^{-1} \tau_0| \\ &= |\gamma_0 \tau + \delta_0| (\lambda + \lambda^{-1} |\tau_0|) |\lambda_1 S_0 \tau + (1 - \lambda_1) e^{i\varphi}| \\ &= |\gamma \tau + \delta| (|S \tau| + |\tau_0|) |\lambda_1 S_0 \tau + (1 - \lambda_1) e^{i\varphi}| \\ &= (|\alpha \tau + \beta| + |\tau_0| |\gamma \tau + \delta|) |\lambda_1 S_0 \tau + (1 - \lambda_1) e^{i\varphi}| \\ &\geq C_1 C' (|\alpha i + \beta| + |\tau_0| |\gamma i + \delta|) \geq C_1 C' |\gamma i + \delta| |S i + \tau_0|. \end{aligned}$$

wobei C' das Minimum von $|\lambda_1 \tau_1 + (1 - \lambda_1) e^{t\varphi}|$ auf der abgeschlossenen Mannigfaltigkeit

$$|\tau_1| = 1, \quad \Im \tau_1 \geq 0, \quad 0 \leq \lambda_1 \leq 1.$$

$C' = C'(\tau_0)$ ist offenbar positiv, da $\lambda_1 \tau_1 + (1 - \lambda_1) e^{t\varphi} = 0$ mit $\Im \tau_0 > 0$ nicht verträglich ist. Hilfssatz 2 ist also mit $C_2 = C_1 C'$ bewiesen.

Die absolute und gleichmäßige Konvergenz von $\Omega(\tau, \tau, A, G)$ und damit von $\Xi_{-r}(\tau, v, A, G, (q))$ im Bereich (55) ergibt sich nach (48) auf Grund der bewiesenen Hilfssätze leicht aus folgender Abschätzung:

$$B = AS = \begin{pmatrix} b_1 & b_2 \\ b_1' & b_2' \end{pmatrix}, \quad A = (a_1', a_2'), \quad S = (\gamma, \delta), \quad \frac{a_2^{(v)'}}{a_1^{(v)'}} = -\frac{\overline{\tau_0^{(v)}}}{\tau_0^{(v)}} \text{ für } v > t,$$

$$\begin{aligned} |h_B(\tau)| &= \prod_{v=1}^t |b_1^{(v)'} \tau^{(v)} + b_2^{(v)'}|^{r/2} |b_1^{(v)} \tau^{(v)} + b_2^{(v)}|^{r/2} \prod_{v=t+1}^n |a_1^{(v)'}|^r |\gamma^{(v)} \tau^{(v)} + \delta^{(v)}|^r \\ &\geq C_1^t C_2^{r(n-t)} |h_B(i)|. \end{aligned}$$

Die Poincaréschen Reihen $\Xi_{-r}(\tau)$ sind daher in § regulär. Die Transformationsformel (41) besagt, daß es sich um automorphe Formen vom Typus $\{G, -r, v\}$ handelt. Es bleibt nur noch das Verhalten von $\Xi_{-r}(\tau)$ in den parabolischen Spitzen von G zu untersuchen. Für Untergruppen G der Hilbertschen Modulgruppe ist $\infty = \{\infty, \infty, \dots, \infty\}$ parabolische Spitze, und es gibt einen Fundamentalbereich \mathfrak{F} zu G [s. 1)] mit folgender Eigenschaft. Alle Punkte $\tau = x + iy$ von \mathfrak{F} , für welche $Ny \geq \xi$ mit hinreichend großem ξ , sind enthalten in einer Punktmenge der Art

$$(56) \quad C' \sqrt[n]{Ny} \leq y^{(v)} \leq C'' \sqrt[n]{Ny}, \quad Ny \geq \xi, \quad |x^{(v)}| \leq C''' \text{ für } v = 1, 2, \dots, n.$$

Diese wiederum kann in einen Bereich (55) eingebettet werden. $\Xi_{-r}(\tau)$ ist also in (56) absolut und gleichmäßig konvergent. Nähert man sich innerhalb \mathfrak{F} der parabolischen Spitze ∞ , so geht, wie man leicht sieht, jedes Glied der Reihe (35) gegen 0, infolgedessen konvergiert $\Xi_{-r}(\tau)$ selbst gegen 0. $\Xi_{-r}(\tau)$ ist also in der Spitze ∞ regulär und verschwindet. Derselbe Sachverhalt trifft für alle parabolischen Spitzen von G zu. Der Beweis wird wie üblich so geführt, daß man eine vorgegebene parabolische Spitze nach ∞ transformiert und die entsprechend transformierte Poincarésche Reihe betrachtet. Das Ergebnis dieses Paragraphen fassen wir zusammen in

Satz 1. Die Poincaréschen Reihen $\Xi_{-r}(\tau, v, A, G, (q))$ sind für $r > 2$, $|v| = 1$ und jedes ganzzahlige Exponentensystem (q) mit $q_n \geq 0$ für $\mu > t$ in ganz \mathfrak{F} absolut konvergent und stellen Spitzenformen vom Typus $\{G, -r, v\}$ dar.

§ 3.

Anwendung der Metrisierungstheorie auf die Poincaréschen Reihen.

Es seien τ, s reelle Zahlen, S eine reelle unimodulare Substitution, $f(\tau)$ und $\varphi(\tau)$ Funktionen über \mathfrak{H} . Wir erklären im Einklang mit der Petersson'schen Art der Bezeichnung

$$(57) \quad \begin{aligned} f \Big|_{(r,s)} S &= f(\tau) \Big|_{(r,s)} S = f(S\tau) N(\gamma\tau + \delta)^{-r} N|\gamma\tau + \delta|^{-s}, \\ f(\tau) \Big|_{(r,s)} S &= f(\tau) \Big|_{(r,s)} S, \quad f(\tau) \Big|_{(s)} S = f(\tau) \Big|_{(0,s)} S, \\ W_r(f(\tau), \varphi(\tau); \mathfrak{B}) &= \int_{\mathfrak{B}} \dots \int_{\mathfrak{B}} f(\tau) \overline{\varphi(\tau)} N(y\tau^{-2} dx dy) \quad (\mathfrak{B} \subset \mathfrak{H}, \tau = x + iy). \end{aligned}$$

Dann gilt für reelle unimodulare Substitutionen S_1, S_2

$$(57a) \quad f(\tau) \Big|_{(r,s)} S_1 S_2 = \sigma^{(r)}(S_1, S_2) (f(\tau) \Big|_{(r,s)} S_1) \Big|_{(r,s)} S_2.$$

Der Übergang von den Integrationsvariablen τ zu $S\tau$ liefert die allgemeine Transformationsformel

$$(58) \quad W_{r+s}(f \Big|_{(r,s)} S, \varphi \Big|_{(r,s)} S; S^{-1}\mathfrak{B}) = W_{r+s}(f, \varphi; \mathfrak{B}),$$

insbesondere also

$$(59) \quad W_r(f \Big|_{(r)} S, \varphi \Big|_{(r)} S; S^{-1}\mathfrak{B}) = W_r(f \Big|_{(r)} S, \varphi \Big|_{(r)} S; S^{-1}\mathfrak{B}).$$

Für automorphe Formen $f(\tau), \varphi(\tau)$ vom Typus $\{\mathbf{G}, -r, v\}$ ist ferner

$$(60) \quad f^S(\tau) = f(\tau) \Big|_{(r)} S, \quad |f^S(\tau)| = |f(\tau)| \Big|_{(r)} S,$$

mithin

$$(61) \quad W_r(f^S, \varphi^S; S^{-1}\mathfrak{B}) = W_r(f, \varphi; \mathfrak{B})$$

sowie

$$(62) \quad W_r(f, \varphi; L^{-1}\mathfrak{B}) = W_r(f, \varphi; \mathfrak{B}) \quad \text{für } L \in \mathbf{G}, v\bar{v} = 1;$$

denn in diesem Fall gilt $f^L = v(L)f$, $\varphi^L = v(L)\varphi$. Bezeichnet \mathfrak{F} einen Fundamentalbereich für \mathbf{G} in \mathfrak{H} , so ist das skalare Produkt von f und φ erklärt durch

$$(63) \quad (f, \varphi) = (f, \varphi)_{\mathbf{G}} = W_r(f, \varphi; \mathfrak{F}).$$

Es existiert nach ¹⁾, wenn $f \cdot \varphi$ eine Spitzenform für \mathbf{G} ist, und hat die Eigenschaft

$$(64) \quad (f, \varphi)_{\mathbf{G}} = (f^S, \varphi^S)_{S^{-1}\mathbf{G}S}$$

für beliebige reelle unimodulare Transformationen S .

Wir berechnen jetzt das Skalarprodukt einer gegebenen Spitzenform $f(\tau)$ vom Typus $\{\mathbf{G}, -r, v\}$ mit der Poincaréschen Reihe $\Xi_{-r}(\tau, v, A, \mathbf{G}, (q))$,

$r > 2$ und $|v| = 1$ vorausgesetzt. Bei der Durchführung sind mehrmals Vertauschungen von Grenzprozessen vorzunehmen. Um einzusehen, daß solche Vertauschungen statthaft sind, wählen wir nach 3) einen Fundamentalbereich \mathfrak{F} für G mit der folgenden Eigenschaft. Es sei s_1, s_2, \dots, s_n ein vollständiges System von parabolischen Spitzen von G , die paarweise nach G nicht äquivalent sind. Die reelle unimodulare Transformation Q_μ möge s_μ nach ∞ transformieren: $Q_\mu s_\mu = \infty$. Zu jeder parabolischen Spitze s_μ gibt es nun eine Punktmenge \mathfrak{B}_μ , so daß $Q_\mu \mathfrak{B}_\mu$ in dem Bereich

$$(65) \quad C' \sqrt[n]{Ny} \leq y^{(r)} \leq C'' \sqrt[n]{Ny}, \quad Ny \geq \xi_0, \quad |x^{(r)}| \leq C''' \quad \text{für } r = 1, 2, \dots, n$$

mit gewissen positiven Konstanten C', C'', C''', ξ_0 enthalten ist und

$$(66) \quad \mathfrak{F} = \sum_{\mu=1}^n \mathfrak{B}_\mu$$

einen Fundamentalbereich für G darstellt. Wir nehmen die Bereiche \mathfrak{B}_μ paarweise punktfremd an. Entsprechend der Zerlegung (66) gilt dann

$$(67) \quad (f, \Xi_{-r}) = \sum_{\mu=1}^n W_r(f, \Xi_{-r}; \mathfrak{B}_\mu).$$

Nach (46), (64) genügt es, dieses Skalarprodukt für den Fall zu berechnen, daß s die spezielle Lage (44) hat und $A = A_t$ ist. Es handelt sich dann um die Poincarésche Reihe (47), die wir auch in der Form

$$(68) \quad \Xi_{-r}(\tau, v, A_t, G, (q)) = \sum_{s \in \mathfrak{S}_s(G)} \bar{v}(S) \vartheta_{(q)}(\tau) |S|$$

mit

$$(69) \quad \vartheta_{(q)}(\tau) = (h_{A_t}(\tau))^{-1} \prod_{r=1}^t (\tau^{(r)2} \pi^{t(r)}) \prod_{r=t+1}^n \left(\frac{\tau^{(r)} - i}{\tau^{(r)} + i} \right)^{q_r}$$

schreiben können. Sei C_0 eine Schranke für den Betrag von $\vartheta_{(q)}(\tau) h_{A_t}(\tau)$ [bei festem (q)], dann hat $\frac{1}{C_0} \Xi_{-r}(\tau)$ die Majorante

$$\Omega(\tau, r, A_t, G) = \sum_{s \in \mathfrak{S}_s(G)} |h_{A_t}(\tau)|^{-1} |S|.$$

Sie konvergiert, wie wir sahen, für alle $\tau \in \mathfrak{H}$. Für eine beliebige Teilmenge $\mathfrak{I} \subset \mathfrak{S}_s(G)$ bilden wir die Teilreihe

$$R(\tau, \mathfrak{I}) = \sum_{s \in \mathfrak{I}} |h_{A_t}(\tau)|^{-1} |S|$$

und

$$V_\mu(\mathfrak{I}) = W_r(|f(\tau)|, R(\tau, \mathfrak{I}); \mathfrak{B}_\mu).$$

Auf Grund der Formeln (59), (60) und der Beziehung

$$R(\tau, \mathfrak{I}) \Big|_L = R(\tau, \mathfrak{I}L) = N|c\tau + d|^{-r} R(L\tau, \mathfrak{I}), \quad \underline{L} = (c, d),$$

für eine reelle unimodulare Substitution L gilt dann

$$V_{\mu}(\mathfrak{T}) = W_r(|f^{Q_{\mu}^{-1}}(\tau)|, R(\tau, \mathfrak{T} Q_{\mu}^{-1}); Q_{\mu} \mathfrak{B}_{\mu}).$$

Wie wir oben sahen, konvergiert $R(\tau, \mathfrak{T} Q_{\mu}^{-1})$ in $Q_{\mu} \mathfrak{B}_{\mu}$ gleichmäßig und es gibt eine positive Konstante C'_{μ} , so daß

$$R(\tau, \mathfrak{T} Q_{\mu}^{-1}) \leq C'_{\mu} R(i, \mathfrak{T} Q_{\mu}^{-1}) = C'_{\mu} N | -c_{\mu} i + a_{\mu} |^{-r} R(Q_{\mu}^{-1} i, \mathfrak{T})$$

mit $Q_{\mu}^{-1} = (-c_{\mu}, a_{\mu})$

und damit

$$V_{\mu}(\mathfrak{T}) \leq C''_{\mu} R(Q_{\mu}^{-1} i, \mathfrak{T}) \int \dots \int_{Q_{\mu} \mathfrak{B}_{\mu}} |f^{Q_{\mu}^{-1}}(\tau)| N(y^{r-2} dx dy).$$

Gleichmäßig für alle Systeme \mathfrak{T} ist also

$$0 \leq V_{\mu}(\mathfrak{T}) \leq C_{\mu} R(Q_{\mu}^{-1} i, \mathfrak{T}) \quad \text{für } \mu = 1, 2, \dots, h$$

mit konstantem C_{μ} . Aus diesen Abschätzungen ergeben sich die Existenz von $W_r(f(\tau), \bar{v}(S) \vartheta_{(q)}(\tau) | S; \mathfrak{F})$ für $S \in \mathfrak{S}_s(\mathbf{G})$, die Konvergenz von

$$(70) \quad \sum_{S \in \mathfrak{S}_s(\mathbf{G})} W_r(|f(\tau)|, |\vartheta_{(q)}(\tau)| | S; \mathfrak{F})$$

und schließlich die Erlaubnis, Integration und Summation zu vertauschen, wenn man die Reihe (68) in (f, \mathfrak{E}_{-r}) einträgt. Beachtet man noch $f^S(\tau) = v(S) f(\tau)$ für $S \in \mathbf{G}$ und Regel (58), so folgt also

$$(71) \quad (f, \mathfrak{E}_{-r}) = \sum_{S \in \mathfrak{S}_s(\mathbf{G})} W_r(f^S(\tau), \vartheta_{(q)}(\tau) | S; \mathfrak{F}) \\ = W_r(f(\tau), \vartheta_{(q)}(\tau); \mathfrak{F}_0),$$

wobei $\mathfrak{F}_0 = \sum_{S \in \mathfrak{S}_s(\mathbf{G})} S \mathfrak{F}$ einen einfach überdeckten Fundamentalbereich für die Gruppe $\mathbf{A}(s)$ darstellt. Da die Poincaréschen Reihen von der Auswahl des Systems $\mathfrak{S}(A, \mathbf{G})$ nicht abhängen, so schließt man für $S_0 \in \mathbf{A}(s)$, $S \in \mathbf{G}$, daß

$$\bar{v}(S) \vartheta_{(q)}(\tau) | S = \bar{v}(S_0 S) \vartheta_{(q)}(\tau) | S_0 S \\ = \sigma^{(r)}(S_0, S) \bar{v}(S_0 S) (\vartheta_{(q)}(\tau) | S_0) | S,$$

$$(72) \quad \vartheta_{(q)}(\tau) | S = \bar{v}(S_0) (\vartheta_{(q)}(\tau) | S_0) | S, \\ \vartheta_{(q)}(\tau) | S_0 = v(S_0) \vartheta_{(q)}(\tau).$$

D. h. $\vartheta_{(q)}(\tau)$ ist eine automorphe Form vom Typus $\{\mathbf{A}(s), -r, v\}$, und daher darf \mathfrak{F}_0 in (71) durch einen beliebigen anderen Fundamentalbereich \mathfrak{F}^* von $\mathbf{A}(s)$ ersetzt werden. Diese Umformungen sind wegen der Konvergenz von (70) zulässig. Bezeichnen wir wieder mit λ_{μ}^2 die Multiplikatoren der Erzeugenden S_{μ} von $\mathbf{A}(s)$ und berechnen z_1, z_2, \dots, z_t aus dem Gleichungssystem (18) [darin ist $\tau_1^{(v)} = \tau^{(v)}$ für $v = 1, 2, \dots, t$], so hat man in

$$(73) \quad -\frac{1}{2} \leq \operatorname{Re} z_{\mu} \leq \frac{1}{2} \quad \text{für } \mu = 1, 2, \dots, t$$

einen geeigneten Fundamentalbereich \mathfrak{F}^* . Es sei α, β, λ ein System positiver Zahlen und $\mathfrak{B}_{\alpha\beta\lambda}$ definiert als der Durchschnitt von \mathfrak{F}^* mit der Punktmenge

$$(74) \quad \begin{aligned} \alpha^{(v)} \leq \arg \tau^{(v)} \leq 1 - \beta^{(v)} & \quad \text{für } v \leq t, \\ \left| \frac{\tau^{(v)} - i}{\tau^{(v)} + i} \right| \leq 1 - \lambda^{(v)} & \quad \text{für } v > t. \end{aligned}$$

$\mathfrak{B}_{\alpha\beta\lambda}$ ist eine abgeschlossene Punktmenge im Innern von \mathfrak{H} . Die bisherigen Betrachtungen haben dann ergeben, daß

$$(75) \quad (f, \Xi_{-r})_{\mathfrak{G}} = (f, \vartheta_{(q)})_{\mathfrak{A}(s)} = \lim_{\alpha, \beta, \lambda \rightarrow 0} W_r(f(\tau), \vartheta_{(q)}(\tau); \mathfrak{B}_{\alpha\beta\lambda}).$$

Hierin muß jetzt die Entwicklung von $f(\tau)$ zum Punktsystem s eingetragen werden; sie lautet [nach (34) für $A = A_t, S = E$]:

$$(76) \quad f(\tau) = \sum_{k_1, k_2, \dots, k_n} b_{k_1 k_2 \dots k_n} \vartheta_{(k)}(\tau)$$

und ist als Laurent-Reihe in jedem abgeschlossenen Bereich von \mathfrak{H} , insbesondere $\mathfrak{B}_{\alpha\beta\lambda}$, absolut und gleichmäßig konvergent. Daher darf die Integration über $\mathfrak{B}_{\alpha\beta\lambda}$ mit der Summation über (k) vertauscht werden. Wir erhalten dann

$$(77) \quad (f, \Xi_{-r}) = \lim_{\alpha, \beta, \lambda \rightarrow 0} \sum_{(k)} b_{k_1 k_2 \dots k_n} W_r(\vartheta_{(k)}(\tau), \vartheta_{(q)}(\tau); \mathfrak{B}_{\alpha\beta\lambda}).$$

Zur Berechnung von

$$W_r(\vartheta_{(k)}(\tau), \vartheta_{(q)}(\tau); \mathfrak{B}_{\alpha\beta\lambda})$$

führen wir neue Integrationsvariable ein:

$$1. \quad v \leq t. \quad \log \tau = \log(x + iy) = u + iv, \quad x = e^u \cos v, \quad y = e^u \sin v,$$

$$|\tau|^{-r} y^{r-2} \left| \frac{\partial(x, y)}{\partial(u, v)} \right| = \sin^{r-2} v,$$

$$2. \quad v > t. \quad w = \frac{\tau - i}{\tau + i} = u + iv = \varrho e^{i\psi},$$

$$|\tau + i|^{-2r} y^{r-2} \left| \frac{\partial(x, y)}{\partial(u, v)} \right| = 4^{1-r} (1 - \varrho^2)^{r-2}, \quad \left| \frac{\partial(u, v)}{\partial(\varrho, \psi)} \right| = \varrho.$$

Schließlich wird die Variablenreihe $u^{(1)}, u^{(2)}, \dots, u^{(t)}$ noch ersetzt durch die ersten t der n Variablen

$$(78) \quad \sigma_\mu = \Re z_\mu = \sum_{v=1}^t m_\mu^{(v)} u^{(v)} \quad \text{für } \mu = 1, 2, \dots, n$$

mit

$$\left| \frac{\partial(u^{(1)}, u^{(2)}, \dots, u^{(t)})}{\partial(\sigma_1, \sigma_2, \dots, \sigma_t)} \right| = R,$$

wobei R der Betrag der Determinante $|\log \lambda_\mu^{(v)}|$ ($\mu, v \leq t$). In den neuen Veränderlichen wird die Punktmenge $\mathfrak{B}_{\alpha\beta\lambda}$ beschrieben durch die Ungleichungen

$$(79) \quad \left. \begin{aligned} -\frac{1}{2} &\leq \sigma_v \leq \frac{1}{2} \\ \alpha^{(v)} &\leq \psi^{(v)} \leq \pi - \beta^{(v)} \end{aligned} \right\} \text{ für } v = 1, 2, \dots, t,$$

$$\left. \begin{aligned} 0 &\leq \varrho^{(v)} \leq 1 - \lambda^{(v)} \\ 0 &\leq \psi^{(v)} \leq 2\pi \end{aligned} \right\} \text{ für } v = t+1, t+2, \dots, n.$$

Erinnert man sich noch der Beziehungen

$$(80) \quad \left. \begin{aligned} j^{(v)} &= \sum_{\mu=1}^n m_\mu^{(v)} (k_\mu + \varkappa_\mu) \\ p^{(v)} &= \sum_{\mu=1}^n m_\mu^{(v)} (q_\mu + \varkappa_\mu) \end{aligned} \right\} \text{ für } v = 1, 2, \dots, t,$$

so erhält man nach elementarer Umrechnung

$$\begin{aligned} W_r(\vartheta_{(k)}(\tau), \vartheta_{(q)}(\tau); \mathfrak{B}_{\alpha\beta\lambda}) &= \int \dots \int_{\mathfrak{B}_{\alpha\beta\lambda}} \vartheta_{(k)}(\tau) \overline{\vartheta_{(q)}(\tau)} N(y^{r-2} dx dy) \\ &= 4^{(1-r)(n-t)} R \cdot \int \dots \int_{\mathfrak{B}_{\alpha\beta\lambda}} \prod_{v=1}^n e^{2\pi i(k_v - q_v)\psi_v} \prod_{v=1}^t e^{-2\pi i(j^{(v)} + p^{(v)})\psi^{(v)}} \sin^{r-2} \psi^{(v)} \times \\ &\quad \times \prod_{v=t+1}^n \varrho^{(v)k_v + q_v + 1} (1 - \varrho^{(v)2})^{r-2} e^{i(k_v - q_v)\psi^{(v)}} \prod_{v=1}^t d\sigma_v d\psi^{(v)} \prod_{v=t+1}^n d\varrho^{(v)} d\psi^{(v)}. \end{aligned}$$

Dieses Integral ist offenbar nur dann von 0 verschieden, wenn $(k) = (q)$ ist. In diesem Fall wird

$$(81) \quad W_r(\vartheta_{(q)}(\tau), \vartheta_{(q)}(\tau); \mathfrak{B}^*) = 4^{(1-r)(n-t)} \pi^{n-t} R \times$$

$$\times \prod_{v=1}^t \int_0^\pi e^{-4\pi p^{(v)}\psi} \sin^{r-2} \psi d\psi \prod_{v=t+1}^n \int_0^1 \varrho^{q_v} (1 - \varrho)^{r-2} d\varrho.$$

Die hier auftretenden bestimmten Integrale ergeben sich aus den Formeln

$$(82) \quad \int_0^\pi e^{-4\pi p\psi} \sin^{r-2} \psi d\psi = \frac{\pi e^{-2\pi p} \Gamma(r-1)}{2^{r-2} \Gamma\left(\frac{r}{2} + 2\pi i p\right) \Gamma\left(\frac{r}{2} - 2\pi i p\right)} {}^7),$$

$$\int_0^1 \varrho^q (1 - \varrho)^{r-2} d\varrho = \frac{\Gamma(q+1) \Gamma(r-1)}{\Gamma(q+r)}.$$

⁷⁾ E. E. Kummer, De integralibus definitis et seriebus infinitis. Crelle's Journal f. d. r. u. a. Math. 17 (1837), Nr. 11, S. 210–227. Die Formeln (31) und (32) auf S. 217 liefern das gewünschte Integral.

Damit ist die Berechnung des Skalarproduktes durchgeführt mit dem Ergebnis

$$(f(\tau), \Xi_{-\tau}(\tau, v, A, G, (q))) = \varepsilon_r(A_t, (q)) b_{q_1, q_2, \dots, q_n},$$

$$(83) \quad \varepsilon_r(A_t, (q)) = \frac{(4^{(1-r)/2} \pi \Gamma(r-1))^n 2^r R \prod_{r=1}^l e^{-2p^{(r)} \pi^2} \prod_{r=l+1}^n q_r!}{\prod_{r=1}^l \Gamma\left(\frac{r}{2} + 2\pi i p^{(r)}\right) \Gamma\left(\frac{r}{2} - 2\pi i p^{(r)}\right) \prod_{r=l+1}^n \Gamma(r+q_r)}.$$

Die Übertragung dieses Resultats auf den Fall eines beliebig gelegenen Fixpunktsystems s wird durch die Transformationstheorie geleistet und bietet keinerlei Schwierigkeiten mehr. Nach (64) und (46) ist nämlich

$$\begin{aligned} (f(\tau), \Xi_{-\tau}(\tau, v, A, G, (q)))_G \\ = \overline{\zeta^{(v)}(A, L)^{-1} N \lambda^{-\tau} (f^L(\tau), \Xi_{-\tau}(\tau, v^L, A_t, L^{-1} G L, (q)))}_{L^{-1} G L} \end{aligned}$$

mit $L = \lambda A^{-1} A_t$, $\lambda^2 |A^{-1} A_t| = 1$, $\lambda > 0$. Ferner ergibt sich nach (28), (32), (45) aus der Entwicklung

$$(84) \quad h_A(\tau) f(\tau) = \sum_{k_1, k_2, \dots, k_n} b_{k_1, k_2, \dots, k_n}(A) \prod_{r=1}^l \langle A^{(r)} \tau^{(r)} \rangle^{2\pi i p^{(r)}} \prod_{r=l+1}^n \langle A^{(r)} \tau^{(r)} \rangle^{k_r},$$

durch Transformation mit L :

$$f^L(\tau) = \sum_{k_1, k_2, \dots, k_n} b_{k_1, k_2, \dots, k_n}(A) (\zeta^{(v)}(A, L))^{-1} N \lambda^{-\tau} \partial_{(k)}(\tau).$$

Aus (83) folgt daher auch die allgemeine Grundformel

$$(85) \quad (f(\tau), \Xi_{-\tau}(\tau, v, A, G, (q))) = \varepsilon_r(A, (q)) b_{q_1, q_2, \dots, q_n}(A)$$

mit

$$(86) \quad \varepsilon_r(A, (q)) = N \lambda^{-2\tau} \varepsilon_r(A_t, (q))$$

$$= \frac{(2^{2-r} \pi \Gamma(r-1))^n R \prod_{r=1}^l e^{-2p^{(r)} \pi^2} \prod_{r=l+1}^n q_r!}{|N|A|^r \prod_{r=1}^l \Gamma\left(\frac{r}{2} + 2\pi i p^{(r)}\right) \Gamma\left(\frac{r}{2} - 2\pi i p^{(r)}\right) \prod_{r=l+1}^n \Gamma(r+q_r)}.$$

Das Skalarprodukt einer beliebigen (ganzen) Form $f(\tau)$ vom Typus $\{G, -\tau, v\}$ mit einer Poincaréschen Reihe $\Xi_{-\tau}(\tau)$ wird in der Weise bestimmt, daß man $f(\tau)$ mit Hilfe der Eisensteinreihen auf eine Spitzenform reduziert. Die Eisensteinreihen geben zum Skalarprodukt keinen Beitrag, weil sie auf allen Spitzenformen im Sinne der Metrisierungstheorie senkrecht stehen (vgl. ¹⁾) und die hier betrachteten Poincaréschen Reihen sämtlich Spitzenformen darstellen.

Die Folgerungen, die wir nun aus der Grundformel (85) ziehen, bilden den Inhalt der nachstehenden beiden Sätze.

Satz 2. Es sei $r > 2$, $|v| = 1$, (q) ein Exponentensystem mit $q_\mu \geq 0$ für $\mu > 1$. Die Poincarésche Reihe $\Xi_{-r}(\tau, v, A, \mathbf{G}, (q))$ zu einem Punktsystem s ist eine Spitzenform vom Typus $\{\mathbf{G}, -r, v\}$. Sie steht auf allen Spitzenformen $f(\tau)$ vom Typus $\{\mathbf{G}, -r, v\}$ senkrecht, in deren Entwicklung (84) zum Punktsystem s der Koeffizient $b_{(q)}(A)$ zum Exponentensystem (q) verschwindet. Durch diese Eigenschaft ist die Poincarésche Reihe bis auf einen konstanten Faktor eindeutig bestimmt. $\Xi_{-r}(\tau)$ verschwindet dann und nur dann identisch in τ , wenn in der Entwicklung (84) der Koeffizient $b_{(q)}(A)$ für sämtliche Spitzenformen $f(\tau)$ vom Typus $\{\mathbf{G}, -r, v\}$ gleich Null ist.

Satz 3. Unter den Voraussetzungen von Satz 2 kann aus der Menge der Poincaréschen Reihen $\Xi_{-r}(\tau, v, A, \mathbf{G}, (q))$ zu festem A eine Basis für die Schar der Spitzenformen vom Typus $\{\mathbf{G}, -r, v\}$ ausgewählt werden (Vollständigkeitsatz).

Eine ausführlichere Begründung dieser Sätze erübrigt sich; denn die Beweise für die analogen Aussagen über Poincarésche Reihen zu parabolischen Spitzen [vgl. ¹⁾] lassen sich auf die vorliegenden Verhältnisse wörtlich übertragen.

(Eingegangen am 2.10.1941.)

Fixpunktklassen.

Teil III.

Mindestzahlen von Fixpunkten.

Von

Franz Wecken in Marburg a. d. Lahn.

Einleitung.

Nachdem in Teil I¹⁾ mit Hilfe der *Fixpunkt*klasse und der *algebraischen Fixpunktzahl* eine untere Abschätzung für die geometrische Fixpunktzahl einer Abbildung eines Polyeders \mathfrak{P} in sich gewonnen wurde und in Teil II²⁾ aus den gleichen Begriffen weitere Invarianzaussagen hergeleitet wurden, soll in dieser Mitteilung die Mindestzahl geometrisch verschiedener Fixpunkte für die Abbildungsklassen gewisser Polyeder angegeben werden³⁾.

Das angestrebte Ergebnis $m = \mu$, die Identität der geometrischen Fixpunktmindestzahl mit der für die Abbildungsklasse charakteristischen Zahl der wesentlichen Fixpunktklassen, wird in § 1 für die Klasse der Identität, d. h. für alle zur identischen Abbildung homotopen Abbildungen, bewiesen (Satz 2). An einem Beispiel wird zunächst gezeigt, daß der Satz nicht allgemein für zusammenhängende Polyeder gültig ist; eine schärfere Voraussetzung über den Zusammenhang von \mathfrak{P} erscheint notwendig. Der hier benutzte Begriff des d -dimensionalen Zusammenhangs (nach R.) ist gleichbedeutend mit der Bedingung, daß kein höchstens $(d - 2)$ -dimensionales Teilpolyeder $\mathfrak{R} \subset \mathfrak{P}$ existiert, so daß $\mathfrak{P} - \mathfrak{R}$ unzusammenhängend wird. Die weiteren Sätze gelten für zweidimensional zusammenhängende Polyeder. Die gewünschte Abbildungsfunktion auf \mathfrak{P} wird zunächst auf den nulldimensionalen Zellen einer Zerlegung definiert und schrittweise von Dimension zu Dimension aufsteigend über alle Zellen fortgesetzt. Hierzu bietet der Hilfsatz 3 die Handhabe, indem er die Funktion vom Rande einer Zelle auf ihr Inneres fortzusetzen gestattet. Dabei ist der Umstand wichtig, daß die auf

¹⁾ Fixpunktklassen I. Math. Annalen 117 (1941), S. 659–671 (zit. als W.). Dort sind in § 1 die häufig benutzten Bezeichnungen eingeführt. — Weitere Literatur: H. Hopf, Vektorfelder in n -dimensionalen Mannigfaltigkeiten. Math. Annalen 96 (1927), S. 225–250 (zit. als H.); P. Alexandroff und H. Hopf, Topologie I. Berlin 1935 (zit. als AH.); K. Reidemeister, Topologie der Polyeder und kombinatorische Topologie der Komplexe. Leipzig 1938 (zit. als R.).

²⁾ Fixpunktklassen II. Math. Annalen 118 (1941), S. 216. Diese Arbeit wird hier nicht benutzt.

³⁾ Vgl. den Vorbericht in W., S. 661f.

der Begrenzung \mathfrak{A} eines Teilbereiches \mathfrak{A} von \mathfrak{P} definierte Abbildung die bei Fortsetzung über \mathfrak{A} auftretende Fixpunktzahl $v(g, \mathfrak{A})$ bereits festlegt (Hilfssatz 4); dies ermöglicht die Definition der „Grenzfixpunktzahl“ $v_G(g, \mathfrak{A}) = v(g, \mathfrak{A})$. Bei der Fortsetzung über solche Nebensimplexe, die das Gebiet, auf dem die Abbildung noch nicht definiert ist, zerlegen, ist also dafür zu sorgen, daß die entstehenden Teilgebiete bis auf höchstens eins die Grenzfixpunktzahl Null erhalten (Hilfssatz 9). Die Existenz der gewünschten Abbildung (Satz 2) ergibt sich schließlich aus dem schärferen Satz 1, der eine auf einem Teilpolyeder \mathfrak{A} gegebene Abbildung über das Gebiet $\mathfrak{P} - \mathfrak{A}$ fortzusetzen ermöglicht.

Das Ergebnis von § 1 wurde für differenzierbare Mannigfaltigkeiten (mit einer zusätzlichen Eigenschaft \mathfrak{R}) von Alexandroff und Hopf (AH. XIV, § 4) hergeleitet. Für beliebige Mannigfaltigkeiten ist es in der in Anm. ¹⁾ zitierten Arbeit von H. Hopf enthalten (H., S. 246, Satz IV). Schlußweisen und Begriffsbildungen aus H. sind zum Teil in § 1 verwertet [vgl. Anm. ⁶⁾, ¹³⁾], jedoch ist der dort grundlegende Begriff des komplexstetigen Vektorfeldes hier vermieden bzw. durch den Begriff der Abbildung ersetzt, der dem Verf. geeigneter erscheint ⁴⁾. Vektorfelder im Polyeder werden nur unter stark einschränkenden Voraussetzungen verwendet. Neu sind weiter die Überlegungen, die durch das Hinausgehen über Mannigfaltigkeiten erforderlich werden, besonders in den Beweisen der Hilfssätze 1, 5, 6 und 8.

In § 2 wird die Gültigkeit von $m = \mu$ auf alle Abbildungsklassen ausgedehnt; der hier durchgeführte Beweis erfordert eine schärfere Voraussetzung über \mathfrak{P} als die in § 1 gemachte, es wird sogar etwas mehr als dreidimensionaler Zusammenhang vorausgesetzt. Der Beweis geht von einer Abbildung mit nur regulären Fixpunkten aus und ersetzt zunächst zwei zur gleichen Fixpunktklasse gehörige, p und q , durch einen einzigen. Nach der in W., S. 665, angegebenen Definition der Fixpunktklassen gibt es eine Kurve \mathfrak{C} von p nach q , die zu ihrem Bilde $f(\mathfrak{C}) = \mathfrak{C}'$ homotop ist. Diese Kurve einschließlich ihrer Endpunkte wird mit einem schlauchförmigen Gebiet umgeben, und nur in diesem Gebiete wird die Funktion $f(x)$ abgeändert. Die Deformation wird in zwei Schritten durchgeführt. Zunächst wird das Bild \mathfrak{C}' von \mathfrak{C} längs einem zwischen \mathfrak{C} und \mathfrak{C}' eingespannten Flächenstück nach \mathfrak{C} deformiert (Hilfssatz 14); man erhält eine Abbildung $g_1(x)$, die \mathfrak{C} punktweise fest läßt, bei der aber \mathfrak{C} und die anderen etwa neu entstandenen Fixpunkte ganz in einem Gebiet liegen, in dem $g_1(x)$ eine „kleine Deformation“ ist. Aus $g_1(x)$

⁴⁾ Die in H., S. 234f. gegebene Definition des Vektors von P nach $f(P)$ konnte sich Verfasser nicht zu eigen machen, da die durch die Umgebungsdarstellung $\Omega_{\mu^n}^n \longleftrightarrow E_{\mu^n}^n$ erklärte Geradlinigkeit in $\Omega_{\mu^n}^n$ von μ^n abhängt und im Durchschnitt $\Omega_{\mu_1^n}^n \cap \Omega_{\mu_2^n}^n$ zweier Umgebungen mehrdeutig werden kann.

erhält man dann durch stetige Abänderung in dem erwähnten Gebiet unter Anwendung von Satz 1 die gewünschte Funktion $g(x)$ (Hilfssatz 15). Entsteht hierbei ein Fixpunkt der Vielfachheit Null, so kann man ihn nachträglich leicht beseitigen. Hat dann die Funktion $g(x)$ noch unwesentliche Fixpunktklassen oder wesentliche mit mehr als einem Fixpunkt, so läßt sich das Verfahren wiederholen und dadurch die geometrische Fixpunktzahl solange vermindern, bis sie gleich der Zahl μ der wesentlichen Fixpunktklassen der neu-gewonnenen Funktion ist. Diese Zahl μ ist nach W. dieselbe wie bei der Abbildung $f(x)$; beim Beweis, daß $m = \mu$ ist, wird also die Deformationsinvarianz von μ explizit benutzt.

Die Arbeit macht über das zugrunde gelegte Polyeder keine anderen als die erwähnten Zusammenhangsvoraussetzungen; insbesondere werden für Mannigfaltigkeiten oder Pseudomannigfaltigkeiten keine weitergehenden Aussagen als für sonstige Polyeder gewonnen. Jedoch fallen, wenn \mathfrak{P} eine Mannigfaltigkeit ist, Teile des Beweisganges fort, so Hilfssatz 1 und 5 und größtenteils die Beweise von Hilfssatz 6 und 8. Diese Teile sind im Text durch eckige Klammern [] gekennzeichnet. Auf weitere bei Mannigfaltigkeiten eintretende Vereinfachungen wird in Anmerkungen hingewiesen, wobei dieser Spezialfall als „Fall (M)“ bezeichnet ist.

§ 1.

Abbildungsklasse der Identität.

Die zur Identität homotopen Abbildungen (die Deformationen, wie man auch sagt) haben höchstens eine wesentliche Fixpunktklasse, weil die identische Abbildung jeden Punkt als Fixpunkt hat und weil jede zusammenhängende Fixpunktmenge einer einzigen Fixpunktklasse angehört. Die Lefschetzsche Zahl oder algebraische Fixpunktzahl dieser Abbildungsklasse ist gleich der Eulerschen Charakteristik des Polyeders⁶⁾:

$$\Lambda = \chi(\mathfrak{P}) = \sum_r (-1)^r \alpha^r$$

(α^r ist die Zahl der r -dimensionalen Zellen einer beliebigen Simplicialzerlegung von \mathfrak{P}). Die Aufgabe ist also, eine Deformation mit einem einzigen bzw. im Falle $\chi = 0$ ganz ohne Fixpunkt anzugeben. Sie wird im folgenden gelöst für *zweidimensional zusammenhängende* Polyeder. Dabei heißt ein Polyeder *d-dimensional zusammenhängend*, wenn es eine Zerlegung besitzt, die keine Grundsimplexe von geringer Dimension als d enthält und in der je zwei d -dimensionale Zellen durch einen aus $(d-1)$ - und d -dimensionalen Zellen bestehenden Weg verbunden sind (vgl. R., S. 45).

⁶⁾ Vgl. W., S. 664.

Daß diese Aufgabe bei nur eindimensionalem Zusammenhang nicht immer lösbar ist, lehrt folgendes

Beispiel. \mathfrak{P} sei das in Fig. 1 dargestellte eindimensionale Polyeder, also $\chi(\mathfrak{P}) = -3$. Wir nehmen per absurdum an, es sei eine Deformation $f(x)$ von \mathfrak{P} in sich mit nur einem Fixpunkt p gegeben. $\{f_\tau\}$ ($0 \leq \tau \leq 1$) sei ein Deformationsweg, der von der Identität f_0 nach $f_1 = f$ führt. f_τ kann, wie man sich überlegt, so normiert werden, daß für jedes $x \in \mathfrak{P}$ die Kurve $\{f_\tau(x)\}$ möglichst kurz wird; es wird dann durch ein Anfangsstück $0 \leq \tau < \varepsilon = \varepsilon(x)$ für $x \neq p$ eine Richtung im Punkte x festgelegt, also ein Richtungsfeld auf \mathfrak{P} erklärt, das bis auf die Eckpunkte und den Fixpunkt stetig ist. Auf jeder Strecke von \mathfrak{P} ist die Richtung konstant, außer wenn sie den Fixpunkt p trägt. Man erkennt weiter: ein Eckpunkt u , der $k+1$ Strecken berandet, wobei auf k Strecken das Richtungsfeld von u fort und auf 1 Strecken nach u hin gerichtet ist, ist entweder ein Fixpunkt der Vielfachheit $v = 1 - k$ oder im Falle $k = 1$ kein Fixpunkt; ein Fixpunkt im Innern einer Strecke kann dementsprechend nur die Vielfachheiten $v = 0, \pm 1$ haben. Da für \mathfrak{P} nach Fig. 1 $k \leq 3$ ist, ist also ein Fixpunkt p mit der Vielfachheit $v = \chi(\mathfrak{P}) = -3$ nicht möglich. Es treten daher bei jeder Deformation von \mathfrak{P} mindestens zwei Fixpunkte auf.

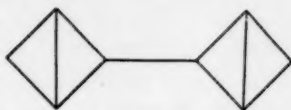


Fig. 1.

In einem eindimensionalen Polyeder sind, wie die Überlegung lehrt, nur Fixpunkte begrenzter Vielfachheit existenzfähig. In jedem r -dimensionalen Simplex ($r > 1$) sind Fixpunkte beliebiger (ganzzahliger) Vielfachheit möglich. Dies ist jedoch nicht das Wesentliche an dem obigen Beispiel; eine nähere Überlegung lehrt, daß man denselben Widerspruch erhält, wenn man \mathfrak{P} dadurch in ein homogen n -dimensionales Polyeder verwandelt, daß man an jede der Strecken ein n -dimensionales Simplex anheftet, ohne daß diese Simplexe anders als an den Ecken zusammenhängen.

Von jetzt ab sei $\mathfrak{P} = |K|$ ein zweidimensional zusammenhängendes Polyeder; insbesondere seien also alle Grundsimplexe von K mindestens zweidimensional. Jeder Punkt x von \mathfrak{P} liegt in genau einem Simplex α von K , dem „Träger“ von x ; wir schreiben $\alpha = Tr(x)$ oder $Tr_K(x)$. Ist \mathfrak{M} eine Punktmenge aus \mathfrak{P} , so sei $\Omega_K(\mathfrak{M})$ die Vereinigungsmenge der Träger der Punkte von \mathfrak{M} und der von den Trägern berandeten Simplexe („offener Stern“; AH., S. 131). Weiter sei für $x \in \mathfrak{P}$ $\mathfrak{B}_K(x)$ die Vereinigungsmenge der Simplexe aus K , die mit $Tr(x)$ mindestens eine Ecke gemeinsam haben⁶⁾. — Aus R. übernehmen wir den Begriff des *Verbindungsproduktes von Polyedern* (S. 82)

⁶⁾ In H., S. 233 als „Simplexumgebung“ Ω_K^n bezeichnet.

und des *Verbindungspolynoms* (mod 2) eines *Simplizialkomplexes* (S. 49). Ein nur aus einem Punkt q bestehendes Polyeder werde mit $|q|$ bezeichnet. — Unter einer *Sphäre* verstehen wir (nach AH., S. 358) ein zu einem Simplex-rande homöomorphes Polyeder. — Für weitere Bezeichnungen vgl. W., § 1.

[Hilfssatz 17]. Ist $\mathfrak{Y} = \Omega \mathfrak{S}^{-1}$ das *Verbindungsprodukt* irgendeines Polyeders Ω mit einer $(r-1)$ -dimensionalen Sphäre \mathfrak{S}^{-1} , so läßt sich (nach Tilgung einer etwa vorhandenen Zerlegung von \mathfrak{Y}) eine krumme *Simplizialzerlegung* K' von \mathfrak{Y} angeben, bei der alle höchstens $(r-1)$ -dimensionalen Simplexe zu Ω *punktfremd* sind.

Beweis. Das Verbindungsprodukt $\Omega \mathfrak{S}$ kann aufgefaßt werden als die Menge der Tripel $\{q, s, \sigma\}$ mit $q \in \Omega$, $s \in \mathfrak{S}$, $0 \leq \sigma \leq 1$, wobei nachträglich für jedes feste $s \in \mathfrak{S}$ alle Punkte $\{q, s, 0\}$ ($q \in \Omega$) sowie für jedes feste $q \in \Omega$ alle Punkte $\{q, s, 1\}$ ($s \in \mathfrak{S}$) identifiziert^{a)} worden sind. Durch $\sigma = 0$ wird \mathfrak{S} , durch $\sigma = 1$ wird Ω bezeichnet. Der durch $\sigma \geq \frac{1}{2}$ bestimmte Teil \mathfrak{Y}_1 von \mathfrak{Y} — nur um ihn brauchen wir uns zu kümmern — ist homöomorph zu einem topologischen Produkt^{b)} $\Omega \times \mathfrak{I}$: \mathfrak{I} ist die Menge der Paare $\{s, \sigma\}$ mit $s \in \mathfrak{S}$, $\frac{1}{2} \leq \sigma \leq 1$, wobei alle Paare $\{s, 1\}$ identifiziert sind; \mathfrak{I} ist also homöomorph zu einem Verbindungsprodukt $|q_0| \mathfrak{S}$, wo q_0 ein geeignet gewählter Punkt ist. Da das Verbindungsprodukt eines Punktes mit einer Sphäre $\mathfrak{S} = \mathfrak{S}^{r-1}$ zu einem abgeschlossenen Simplex α^r homöomorph oder, wie man sagt, ein r -dimensionales Element ist, ist auch \mathfrak{I} ein r -dimensionales Element. Die gewünschte *Simplizialzerlegung* ist also für ein Polyeder vom Typus

$$(1) \quad \mathfrak{X} = \Omega \times \alpha^r$$

anzugeben. Ω liege in einem euklidischen Raum \mathfrak{R}^n mit den Koordinaten ξ_1, \dots, ξ_n ; das Simplex α^r sei in \mathfrak{R}^r (mit den Koordinaten $\xi_{n+1}, \dots, \xi_{n+r}$) realisiert, so daß der Punkt q_0 mit

$$(2) \quad \xi_{n+1} = \dots = \xi_{n+r} = 0$$

im Innern von α^r liegt. \mathfrak{R}^n und \mathfrak{R}^r denken wir in einem Produktraum \mathfrak{R}^{n+r} (mit den Koordinaten ξ_1, \dots, ξ_{n+r}) eingebettet. Ordnet man jedem Punktepaar $\{q, t\}$ ($q \in \Omega$, $t \in \mathfrak{I}$) den Punkt $x \in \mathfrak{R}^{n+r}$ zu, dessen Koordinaten die von q und t sind, so erhält man das topologische Produkt (1) als euklidisches Polyeder \mathfrak{X} in \mathfrak{R}^{n+r} . Ω ist als Teilmenge von \mathfrak{X} durch die Gleichungen (2) gegeben. Wir haben eine Zerlegung von \mathfrak{X} so zu bestimmen, daß die von ihren Eckpunkten u_1, u_2, \dots, u_w aufgespannten abgeschlossenen $(r-1)$ -dimensionalen Simplexe fremd zu Ω sind. $u_{i,k}$ ($k = 1, \dots, n+r$) seien

^{a)} Entfällt im Fall (M); vgl. Einl., S. 3.

^{b)} Vgl. AH., S. 64.

^{c)} AH., S. 35.

die Koordinaten von u_i ($i = 1, \dots, \omega$); wir bestimmen ε so, daß der Abstand von q_0 und a^r größer als 3ε ist, und wählen zunächst eine beliebige ε -Zerlegung¹⁰⁾ von X . Die Ecken u_i , deren Abstandsquadrat von \mathfrak{R}^n

$$\sum_{k=n+1}^{n+r} (u_{ik})^2 < \varepsilon^2$$

ist, sind allein zu berücksichtigen; es seien dies die Punkte u_1, \dots, u_γ . Sie sind, da ihre Projektion auf \mathfrak{R}^r im Innern von a^r liegt, nach der $(n+1)$ -ten bis $(n+r)$ -ten Dimension in gewissen Grenzen frei verschiebbar. Das nützen wir aus, um zu erreichen, daß jede aus der Matrix

$$(u_{ik}) \quad (i = 1, \dots, \gamma; k = n+1, \dots, n+r)$$

gebildete r -reihige Determinante von Null verschieden ist. Damit sind sogar die von je r derartigen Ecken aufgespannten linearen Räume fremd zu dem durch $\xi_{n+1} = \dots = \xi_{n+r} = 0$ bestimmten Raume \mathfrak{R}^n , also zu Ω , und der Satz ist bewiesen.]

Hilfssatz 2. K_1 sei eine Unterteilung von K , $|K| = \mathfrak{P}$, a_1^r eine Zelle von K_1 , $p \in a_1^r$. $g(x)$ sei eine stetige Abbildung des Randes ∂_1^r von a_1^r in \mathfrak{P} mit

$$(3) \quad g(x) \in \mathfrak{P}_K(x),$$

$$(4) \quad g(x) \neq x.$$

Dann gibt es ein Simplex a_2^r mit $p \in a_2^r \subset a_1^r$ und eine stetige Fortsetzung von g über $\overline{a_1^r - a_2^r}$, so daß für $x \in a_1^r - a_2^r$ Formel (3) und (4) und für $q \in \partial_2^r$

$$(5) \quad g(q) \in \overline{\mathfrak{D}_K(p)}$$

gilt¹¹⁾.

Beweis. Wir wählen p als Koordinatenursprung. Die Ortsvektoren der Punkte $g(x)$, q , u , x , y , z bezeichnen wir durch deutsche Buchstaben $g(x)$, q , u , x , y , z und setzen sie auch als Argument in Ortsfunktionen ein, ohne das Funktionszeichen zu ändern: $g(x) = g(x)$. Sind u_1, u_2, \dots die Ecken von K , so stellt sich jeder Punkt $q \in \mathfrak{P}$ eindeutig dar als

$$(6) \quad q = \sum_i \lambda_i u_i \quad (\lambda_i \geq 0, \sum_i \lambda_i = 1),$$

wenn man fordert, daß $\lambda_i \neq 0$ nur gelten darf, wenn u_i Ecke von $Tr_K(q)$ ist. Jeder Koeffizient λ_i hängt stetig von q ab. Wir setzen speziell $q = g(x)$ ($x \in \partial_1^r$) und summieren rechts in (6) gesondert über die Ecken von $\overline{\mathfrak{D}_K(p)}$; dadurch erhalten wir eine eindeutig bestimmte Darstellung¹²⁾

$$(7) \quad g(x) = (1 - \lambda) y + \lambda z,$$

¹⁰⁾ W., S. 663.

¹¹⁾ Erklärung von \mathfrak{D}_K auf S. 547.

¹²⁾ Im Falle $\lambda = 0$, d. h. $g(x) \in \overline{\mathfrak{D}_K(p)}$, bleibt z unbestimmt.

wo $y \in \overline{\mathfrak{D}_K(p)}$ und $Tr_K(z)$ fremd zu $\overline{\mathfrak{D}_K(p)}$ ist. Wegen $g(x) \in \mathfrak{B}_K(x) \subset \mathfrak{B}_K(p)$ ist $\lambda < 1$ und y wohldefiniert. Wegen $g(x) \neq x$ und der Stetigkeit von g auf der abgeschlossenen Menge $\hat{\alpha}_1^r$ gibt es ein η , $0 < \eta < 1$, so daß

$$(8) \quad g(x) \neq \tau x \quad \text{für} \quad 1 - \eta \leq \tau \leq 1$$

ist. Wir setzen fest

$$(9) \quad g(\tau x) = \left(1 - \lambda \frac{\tau + \eta - 1}{\eta}\right) y + \lambda \frac{\tau + \eta - 1}{\eta} z$$

für $x \in \hat{\alpha}_1^r$, $1 - \eta \leq \tau \leq 1$.

Wird für $q \in \hat{\alpha}_1^r$ eine stetige Ortsfunktion $\tau(q)$ erklärt durch

$$q = \tau(q)x; \quad \tau(q) \geq 0, \quad x \in \hat{\alpha}_1^r,$$

so ist die Menge der q mit $\tau(q) < 1 - \eta$ ein Simplex α_2^r , und es ist $p \in \alpha_2^r \subset \alpha_1^r$, $\tau(p) = 0$. Durch (9) ist $g(q)$ auf $\hat{\alpha}_1^r - \alpha_2^r$ stetig erklärt und hat auf $\hat{\alpha}_1^r$ die vorgeschriebenen Randwerte. Für $q = \tau x$, $\tau < 1$ sind $Tr_K(x)$ und $Tr_K(q)$ identisch oder $Tr_K(x)$ berandet $Tr_K(q)$, also ist $\mathfrak{B}_K(x) \subset \mathfrak{B}_K(q)$. Da nach (9) $g(q) \in \mathfrak{B}_K(x)$ gilt, ist also, wie behauptet, $g(q) \in \mathfrak{B}_K(q)$. Zum Nachweis der Fixpunktfreiheit unterscheiden wir zwei Fälle: Ist $\lambda = 0$, so wird nach (9), (7), (8)

$$g(\tau x) = y = g(x) \neq \tau(x);$$

ist $\lambda \neq 0$, so ist für $\tau > 1 - \eta$ $g(\tau x) \notin \overline{\mathfrak{D}_K(p)}$ und daher $g(\tau x) \neq \tau x$, während für $\tau = 1 - \eta$ wieder $g(\tau x) = y$ ist. In diesem Falle liegt y auf der Begrenzung von $\mathfrak{D}_K(p)$, τx aber in α_1^r , also im Innern von $\mathfrak{D}_K(p)$. Jedenfalls ist also $g(\tau x) \neq \tau x$. Allgemein ist für $\tau = 1 - \eta$, d. h. für $q \in \alpha_2^r$ nach (9)

$$g(q) = y \in \overline{\mathfrak{D}_K(p)},$$

w. z. b. w.

Die in diesem Satz und weiterhin auftretende Bedingung (3) besagt, daß die Abbildung $g(x)$ zur Identität benachbart, also eine „kleine Deformation“ ist¹³⁾. Sie ließe sich auch durch eine Ungleichung $\varrho(x, g(x)) < \varepsilon$ ersetzen¹⁴⁾; vgl. S. 575, (66).

Im folgenden werden wir neben Abbildungen Vektor- und Richtungsfelder betrachten. Eine stetige Abbildung $f(x)$ bestimmt ein stetiges Vektorfeld, indem jedem Punkt x der von x nach $f(x)$ führende Vektor zugeordnet wird; sieht man von der Länge des Vektors ab, so erhält man ein Richtungsfeld (AH., S. 535), das bis auf die Fixpunkte von f definiert und stetig ist. Auch für die hier betrachteten Selbstabbildungen von Polyedern ist diese Bildung sinnvoll, solange das Polyeder in einem euklidischen Raum ein-

¹³⁾ In H., S. 234 sind die dieser Bedingung genügenden Abbildungen als Umgebungs transformationen bezeichnet.

¹⁴⁾ $\varrho(x, y)$ bezeichnet den Abstand von x und y (W., S. 662).

gebettet ist. Hier soll jedoch nur dann einer Abbildung f ihr Vektor- bzw. Richtungsfeld zugeordnet werden, wenn die von x nach $f(x)$ verlaufende Strecke ganz dem Polyeder \mathfrak{P} angehört. Liegt die Strecke xy in \mathfrak{P} , so soll die von x nach y weisende Richtung eine *zulässige Richtung* (bezüglich \mathfrak{P} , an der Stelle x) heißen.

Hilfssatz 3. *Unter den Voraussetzungen von Hilfssatz 2 läßt sich $g(x)$ mit (3) über $\overline{a_1'}$ stetig fortsetzen, und zwar, falls a_1' Nebensimplex¹⁵⁾ ist, ohne Fixpunkt; falls a_1' Grundsimplex ist, entweder ohne Fixpunkt oder mit einem solchen einer Vielfachheit ungleich Null an der Stelle p .*

Beweis. Die in Hilfssatz 2 gewonnene Funktion, für die auf \dot{a}_2' (3), (4), (5) gilt, ist noch über $\overline{a_2'}$ fortzusetzen. Erklärt man zunächst für $0 \leq \lambda \leq 1$, $y \in \dot{a}_2'$

$$g_0(\lambda \eta) = \lambda g(\eta),$$

so hat man eine stetige Fortsetzung von g über $\overline{a_2'}$, die den einzigen Fixpunkt p hat und im übrigen das Gewünschte leistet. Falls also a_1' Grundsimplex ist und p eine Vielfachheit $v \neq 0$ hat, setzen wir für $x \in \dot{a}_2'$ $g(x) = g_0(x)$ und sind fertig. Falls p , im Grundsimplex a_1' , die Vielfachheit Null hat, beseitigen wir diesen Fixpunkt durch Abänderung der Abbildung in seiner Umgebung nach AH. XIV, § 2, 3e. Jetzt sei also a_1' Nebensimplex.

Wegen (5) liegt für $q \in \dot{a}_2'$ die Strecke $qg(q)$ in \mathfrak{P} ; der auf \dot{a}_2' erklärten Abbildung $g = g_0$ ist also ein stetiges Richtungsfeld zugeordnet. Die Aufgabe ist, dieses Richtungsfeld stetig über $\overline{a_2'}$ fortzusetzen. Dazu ist die Struktur der Gesamtheit $\mathfrak{V} = \mathfrak{V}(p)$ aller an einer Stelle p zulässigen Richtungen zu untersuchen. \mathfrak{V} ist, wie sich gleich zeigen wird, ein Polyeder. Ist $p \in \dot{a}_1'$, $q \in \dot{a}_1'$, so lassen sich die in p zulässigen Richtungen sämtlich parallel nach q übertragen, man hat also den gleichen Richtungskörper:

$$\mathfrak{V}(p) = \mathfrak{V}(q) = \mathfrak{V}(a_1').$$

Das Richtungsfeld auf \dot{a}_2' stellt eine stetige Abbildung von \dot{a}_1' in $\mathfrak{V}(a_1')$ dar. Die Aufgabe, es stetig über $\overline{a_2'}$ fortzusetzen, läuft darauf hinaus, die Abbildung von \dot{a}_2' in \mathfrak{V} zu einer solchen von $\overline{a_2'}$ in \mathfrak{V} zu erweitern, ist also (nach AH., S. 502, Hilfssatz II) gleichbedeutend mit der Aufgabe, die gegebene Abbildung der Sphäre \dot{a}_2' in \mathfrak{V} homotop auf einen Punkt zusammenzuziehen.

Im Falle (M) ist \mathfrak{V} die Richtungskugel \mathbb{S}^{d-1} , die von höherer Dimension als \dot{a}_2' ist, da a_1' und folglich \dot{a}_2' Nebensimplex ist. Die Abbildung von \dot{a}_2' in \mathfrak{V} ist deshalb unwesentlich, d. h. sie läßt sich so deformieren, daß die Bildmenge von \dot{a}_2' ein echter Teil von \mathfrak{V} ist, und läßt sich folglich (AH., S. 502. Hilfs-

¹⁵⁾ Vgl. AH., S. 157.

satz III) in einen Punkt von \mathfrak{Y} zusammenziehen. Damit ist im Falle (M) Hilfssatz 3 bewiesen.

[Ist im allgemeinen Fall u_1 eine Ecke von K_1 und $e(u_1)$ das Verbindungspolynom von K_1 , so sei $e(u_1) = u_1 e_1(u_1) + e_2(u_1)$ die Entwicklung von $e(u_1)$ an der Stelle u_1 (R., S. 50). Es ist $\Omega_{K_1}(u_1) = |u_1 e_1(u_1)|$; die zulässigen Richtungen an der Stelle u_1 entsprechen daher eindeutig und stetig den Punkten des Polyeders $|e_1(u_1)|$, dessen Zerlegungskomplex der Ring um u_1 heißt (R., S. 50). — Um den Richtungskörper eines Punktes zu ermitteln, der nicht Ecke von K_1 ist, z. B. des Punktes $p \in \alpha_1^r$, hat man p als neuen Eckpunkt einzuführen. Ist $b(u_1) = u_0 u_1 \dots u_r$ das Verbindungspolynom von α_1^r , und $e_3(u_1)$ dasjenige des Ringes um α_1^r , ist also etwa $e(u_1) = b(u_1) e_3(u_1) + e_4(u_1)$, so ist $\Omega_{K_1}(p) = |b(u_1) e_3(u_1)| = |p b(u_1) e_3(u_1)|^{18}$. Der Ring um p hat also jetzt das Verbindungspolynom $b(u_1) e_3(u_1)$ und der Richtungskörper in p ist

$$(10) \quad \mathfrak{Y}(p) = |b(u_1) e_3(u_1)| = \alpha_1^r |e_3(u_1)|.$$

Da α_1^r Nebensimplex ist, ist das Produkt $e_3(u_1)$ nicht leer und das Polyeder

$$\Omega = |e_3(u_1)|$$

enthält mindestens einen Punkt.

Wir wenden auf das Polyeder (10) den Hilfssatz 1 an, erhalten also eine krumme Zerlegung K' von \mathfrak{Y} , deren Zellen bis zur Dimension $r-1$ punktfremd zu Ω sind.

Von der gegebenen Abbildung der Sphäre $\hat{\alpha}_2^r$ in \mathfrak{Y} gehen wir durch homotope Deformation über zu einer simplizialen Approximation (AH. VIII, § 2; W., S. 663) bezüglich K' , bei der also die Bildmenge fremd zu Ω ist. Zuzufolge der Kegelstruktur von $\mathfrak{Y} = \Omega \cdot \alpha_1^r$ kann man die gewonnene Abbildung von $\hat{\alpha}_2^r$ in $\mathfrak{Y} - \Omega$ durch Projektion längs den Erzeugenden des Kegels zu einer solchen in α_1^r deformieren. Ist weiter u_j ein Eckpunkt von Ω , so gelangt man ähnlich durch Deformation innerhalb des in \mathfrak{Y} enthaltenen Kegels $|u_j b(u_1)|$ zu einer Abbildung von $\hat{\alpha}_2^r$ auf $|u_j|$. Damit ist die Abbildung von $\hat{\alpha}_2^r$ in \mathfrak{Y} auf einen Punkt zusammengezogen und der Satz bewiesen.]

Der Hilfssatz 3 wird uns in Verbindung mit den folgenden Überlegungen in die Lage setzen, die gewünschte Deformation von \mathfrak{B} in sich schrittweise zu konstruieren, indem wir sie zunächst in den Eckpunkten und dann, jeweils zur nächsthöheren Dimension aufsteigend, auf allen Zellen erklären. Die so entstehende Abbildung ist zufolge (3) wirklich eine Deformation. Ist nämlich

$$(11) \quad x = \sum_i \lambda_i u_i, \quad y = g(x) = \sum_i \sigma_i u_i$$

¹⁸⁾ Vgl. R., S. 96.

im Sinne der eindeutigen Darstellung (6), so sei

$$(12) \quad \mathfrak{z} = \mathfrak{h}(x) = \frac{\sum_i \sqrt{\lambda_i} \sigma_i u_i}{\sum_i \sqrt{\lambda_i} \sigma_i};$$

aus (3) folgt, daß $Tr_K(x)$ und $Tr_K(g(x))$ eine Ecke gemeinsam haben, daß also $\sum_i \sqrt{\lambda_i} \sigma_i > 0$ ist und daher $h(x)$ stetig von x abhängt. Man kann nun in trivialer Weise (geradlinig) $g(x)$ in $h(x)$ und $h(x)$ in x deformieren.

Die Formeln (11), (12) ermöglichen einen neuen Beweis des Hilfssatzes 1 aus W. (S. 665) durch explizite Angabe der Funktion $a(x, y, \tau)$ für den Fall, daß $y \in \mathfrak{B}_K(x)$, also daß $\overline{Tr_K(x)} \cap \overline{Tr_K(y)}$ nicht leer ist. Denn in (12) wird ein von x und y stetig abhängender Punkt z mit $Tr_K(z) = \overline{Tr_K(x)} \cap \overline{Tr_K(y)}$ angegeben¹⁷⁾, und als Kurve $\{a(x, y, \tau) \mid (0 \leq \tau \leq 1)\}$ kann der Streckenzug $\overline{xz} + \overline{zy}$ gewählt werden, wobei τ zur Bogenlänge proportional ist. Die Bedingung $a(x, x, \tau) = x$ ist erfüllt, denn aus $x = y$ folgt $x = z$.

Von hier an bis zum Schluß des Paragraphen haben $K, K_1, N_1, \mathfrak{R}, \mathfrak{A}$ und g stets folgende Bedeutung: K ist eine Zerlegung von \mathfrak{P} , K_1 eine Unterteilung von K , N_1 ein Teilkomplex von K_1 , $\mathfrak{R} = |N_1|$ also Teilpolyeder von $\mathfrak{P} = |K|$; \mathfrak{A} ist ein Gebiet in \mathfrak{P} mit $\mathfrak{A} \subset \mathfrak{P} - \mathfrak{R}$, $\mathfrak{A} \subset \mathfrak{R}$, d. h. \mathfrak{A} ist eins der Gebiete, in die die offene Menge $\mathfrak{P} - \mathfrak{R}$ zerfällt. $g(x)$ ist erklärt als stetige Abbildung von \mathfrak{R} in \mathfrak{P} ; es ist für $x \in \mathfrak{R}$

$$(3) \quad g(x) \in \mathfrak{B}_K(x),$$

$$(4) \quad g(x) \neq x.$$

Hilfssatz 4. Wird $g(x)$ irgendwie nach Hilfssatz 3 über \mathfrak{A} fortgesetzt, so hängt die Fixpunktzahl $v(g, \mathfrak{A})$ von der Wahl der Fortsetzung nicht ab.

Beweis. Die offene Menge $\mathfrak{P} - \mathfrak{R}$ zerfällt in endlich viele Gebiete $\mathfrak{G}_1 = \mathfrak{A}, \mathfrak{G}_2, \dots, \mathfrak{G}_n$. Über jedes derselben läßt sich nach Hilfssatz 3 die auf \mathfrak{R} erklärte Funktion fortsetzen. $g_1(x)$ und $g_2(x)$ seien zwei solche Fortsetzungen, und es sei

$$g_3(x) = \begin{cases} g_2(x) & \text{für } x \in \mathfrak{A}, \\ g_1(x) & \text{für } x \in \mathfrak{P} - \mathfrak{A}. \end{cases}$$

g_3 ist stetig, da g_1 und g_2 auf \mathfrak{A} übereinstimmen. Bedingung (3) ist für g_3 gültig, g_3 ist also homotop zur Identität. \mathfrak{A} ist ausschließlich aus Neben-

¹⁷⁾ Durch \cap wird der Durchschnitt bezeichnet (W., S. 663).

simplex von K_1 aufgebaut und daher nach Hilfssatz 3 bei g_1 und g_2 fixpunktfrei; $v(g_i, \mathfrak{A})$ ($i = 1, 2, 3$) ist also sinnvoll. Man hat

$$\begin{aligned} v(g_1, \mathfrak{P} - \mathfrak{A}) &= v(g_3, \mathfrak{P} - \mathfrak{A}), \\ v(g_1, \mathfrak{P}) &= v(g_3, \mathfrak{P}) = \chi(\mathfrak{P})^{18}), \end{aligned}$$

folglich (vgl. W., Hilfssatz 4)

$$v(g_1, \mathfrak{A}) = v(g_3, \mathfrak{A}) = v(g_2, \mathfrak{A}).$$

Definition. Haben \mathfrak{A} , \mathfrak{R} und $g(x)$ die auf S. 553 angegebene Bedeutung, so definieren wir als *Grenzfixpunktzahl* (auf \mathfrak{A} bei g) $v_G(g, \mathfrak{A})$ die bei Fortsetzung von $g(x)$ über \mathfrak{A} mit der Bedingung (3) sich ergebende Fixpunktzahl $v(g, \mathfrak{A})$.

Nach Hilfssatz 3 ist die Fortsetzung stets möglich; nach Hilfssatz 4 ist $v(g, \mathfrak{A})$ eindeutig bestimmt. — Ist speziell \mathfrak{A} ein $(n-1)$ -dimensionaler Zyklus im \mathfrak{R}^n und $v(p)$ für $p \in \mathfrak{A}$ der von p nach $g(p)$ führende Vektor, so ist $v_G(g, \mathfrak{A})$ gleich der *Kroneckerschen Charakteristik* des Vektorfeldes v auf \mathfrak{A} (AH. XII, § 3, 1). Die Grenzfixpunktzahl kann also als Verallgemeinerung der Kroneckerschen Charakteristik eines Vektorfeldes angesehen werden.

Hilfssatz 3 und 4 und die Definition von v_G bleiben gültig, wenn K_1 eine krumme Unterteilung von K ist, sofern sie nur zu einer euklidischen Simplicialunterteilung isomorph ist.

Aus Stetigkeitsbetrachtungen schließt man leicht, daß es unter den Voraussetzungen von Hilfssatz 4 ein $\varepsilon > 0$ gibt, so daß aus $\varrho(g(x), h(x)) < \varepsilon$ für $x \in \mathfrak{R}$ stets folgt: $v_G(g, \mathfrak{A}) = v_G(h, \mathfrak{A})^{19)}$. Hieraus folgt weiter: *Gibt es einen stetigen Deformationsweg $g_\tau(x)$ ($0 \leq \tau \leq 1$, $x \in \mathfrak{A}$), so daß stets $g_\tau(x) \neq x$ und $v_G(g_\tau, \mathfrak{A})$ sinnvoll ist, so ist $v_G(g_0, \mathfrak{A}) = v_G(g_1, \mathfrak{A})$.*

[Hilfssatz 5. $a_0 = a_0^{r-1}$ sei ein Nebensimplex aus K , X der Ring²⁰⁾ um a_0 . Die offene Menge $\mathfrak{D}_K(a_0) - a_0$ zerfalle in die Gebiete $\mathfrak{G}_1, \dots, \mathfrak{G}_n$:

$$(13) \quad \mathfrak{D}_K(a_0) - a_0 = \mathfrak{G}_1 + \dots + \mathfrak{G}_n.$$

Das Polyeder $\mathfrak{X} = |X|$ zerfällt dann in n Komponenten²¹⁾ $\mathfrak{X}_1, \dots, \mathfrak{X}_n$ und es ist bei geeigneter Numerierung

$$(14) \quad \overline{\mathfrak{G}}_i = \bar{a}_0 \cdot \mathfrak{X}_i \quad (i = 1, \dots, n).$$

Es sei $0 < k \leq n$; K_1 entstehe aus K durch Zentralunterteilung²²⁾ aller in \mathfrak{G}_k liegenden Simplexe; an die Stelle von (13) tritt

$$\mathfrak{D}_{K_1}(a_0) - a_0 = \mathfrak{G}_1 + \dots + \mathfrak{G}_{k-1} + \mathfrak{G}'_k + \mathfrak{G}_{k+1} + \dots + \mathfrak{G}_n$$

¹⁸⁾ Nach W., § 1, S. 664.

¹⁹⁾ Vgl. den Satz in AH. XIV, § 2, 3 b.

²⁰⁾ Vgl. S. 552.

²¹⁾ AH., S. 49.

²²⁾ AH., S. 135.

mit $\overline{\mathfrak{G}_k} = \overline{a_0} \cdot \mathfrak{X}_k$. Die Begrenzung von \mathfrak{G}_k in \mathfrak{P} ist

$$\mathfrak{G}_k' = a_0 + \mathfrak{D}$$

mit $\mathfrak{D} = a_0 \mathfrak{X}_k'$; $\mathfrak{G}_k'' = \mathfrak{G}_k - \mathfrak{G}_k' - \mathfrak{G}_k \cap \mathfrak{D}^{17}$ ist $(r+1)$ -dimensional zusammenhängend.

Beweis. \mathfrak{X} ist nicht leer, da a_0^{-1} Nebensimplex ist. Es seien x und y Punkte aus \mathfrak{X} . Die offenen Simplexe xa_0 und ya_0 ²²⁾ liegen in $\mathfrak{D}_K(a_0) - a_0$. Man erkennt leicht: dann und nur dann, wenn x und y der gleichen Komponente von \mathfrak{X} angehören, liegen xa_0 und ya_0 innerhalb $\mathfrak{D}_K(a_0) - a_0$ im gleichen Gebiet. Daraus und aus der Gleichung

$$\overline{\mathfrak{D}_K(a_0)} = \overline{a_0} \mathfrak{X}$$

ergibt sich (14). Die Zahl n ist, wie leicht zu sehen, invariant gegen alle Unterteilungen, die a_0 selbst nicht betreffen. — Offenbar bleiben \mathfrak{G}_l und \mathfrak{X}_l für $l \neq k$ beim Übergang von K zu K_1 unberührt.

\mathfrak{G}_k' ist ein Gebiet in \mathfrak{P} , $\mathfrak{G}_k' \subset \mathfrak{G}_k$. Die Begrenzung \mathfrak{G}_k' von \mathfrak{G}_k besteht aus denjenigen Zellen von $\mathfrak{G}_k = \overline{a_0} \mathfrak{X}_k'$, die nicht zu \mathfrak{G}_k' gehören, d. h. die nicht von a_0 berandet werden. Jede Zelle von \mathfrak{G}_k' ist ein Simplex

$$(15) \quad b' = u_{i_1} \dots u_{i_\beta} q'_{i_1} \dots q'_{i_\gamma},$$

wo die u_i Ecken von a_0 und die q'_i Ecken von \mathfrak{X}_k' sind; es sei

$$(16) \quad u_{i_1} \dots u_{i_\beta} = c, \quad q'_{i_1} \dots q'_{i_\gamma} = d'.$$

Ist hierbei $c = a_0$ und $\gamma > 0$ (d. h. b' nicht leer), so ist $b' \subset \mathfrak{G}_k'$, sonst $b' \subset \mathfrak{G}_k$. Also besteht \mathfrak{G}_k' genau aus a_0 und den Zellen von $a_0 \mathfrak{X}_k' = \mathfrak{D}$.

Es ist noch zu zeigen, daß \mathfrak{G}_k nach Fortlassung des Gebietes \mathfrak{G}_k' und seiner Begrenzung (bezüglich \mathfrak{G}_k) $\mathfrak{D} - a_0$ zusammenhängend bleibt, daß m. a. W. \mathfrak{G}_k durch \mathfrak{D} in zwei Gebiete geteilt wird:

$$(17) \quad \mathfrak{G}_k = \mathfrak{G}_k' + \mathfrak{G}_k \cap \mathfrak{D} + \mathfrak{G}_k''.$$

Dazu untersuchen wir die Struktur von \mathfrak{X}_k genauer.

Die in (15) auftretenden Punkte q'_i , d. h. die Eckpunkte von \mathfrak{X}_k' , sind genau die neu eingeführten Eckpunkte in \mathfrak{G}_k ; denn ein $q'_i \in \mathfrak{X}_k$ kann nicht darunter sein, weil sonst b' die nicht in K_1 enthaltene Zelle $u_{i_1} \dots u_{i_\beta} q'_i$ als Seite hätte. Die Zellen von \mathfrak{X}_k seien als d_1, d_2, \dots durchnummeriert; der im Innern von $\overline{b_i} = \overline{a_0} d_i$ liegende neue Eckpunkt von K_1 heiße q'_i ($i = 1, 2, \dots$). q'_i ist eindeutig darstellbar als

$$(18) \quad q'_i = \lambda_i p_i + (1 - \lambda_i) q_i$$

²²⁾ xa ist der (offene) Kegel mit der Spitze x und der Basis a (R., S. 16).

mit $0 < \lambda_i < 1$, $p_i \in a_0$, $q_i \in b_i$. Die q_i bestimmen eine reguläre Unterteilung X_{k+1} von X_k , und die q'_i eine zu X_{k+1} isomorphe Zerlegung X'_{k+1} von X'_k . Die Zuordnung $q_i \rightarrow q'_i$ bestimmt eine topologische Abbildung von X_k auf X'_k , die durch $x \rightarrow x'$, $b_i \rightarrow b'_i$ usw. bezeichnet werde. Wir zeigen zunächst: es ist

$$(19) \quad \mathcal{D} \cap b_i = \mathcal{D}_i = a_0 \bar{b}_i - a_0 b'_i,$$

und b_i wird durch \mathcal{D} in zwei zusammenhängende Teile zerlegt gemäß

$$(20) \quad b_i = \mathcal{D}'_i + \mathcal{D}_i + \mathcal{D}''_i, \quad \mathcal{D}'_i \subset \mathcal{G}'_k.$$

Stets ist $b'_i \subset b_i$, und da $X'_k = \sum_i b'_i$ ist und die b_i paarweise punktfremd sind folgt $X'_k \cap b_i = b'_i$. Man sieht nun, daß sowohl der links stehende wie der rechts stehende Ausdruck in (19) genau die Menge derjenigen Zellen (15) bezeichnet, für die (mit den Bezeichnungen (16)) $c \subset a_0$, $d' \subset b'_i$ ist.

Der Rand des Simplex b_i ist ²⁴⁾

$$\bar{b}_i = a_0 \bar{b}_i + a_0 b'_i;$$

da X'_k fremd zu $a_0 X_k$ und daher $\mathcal{D} \cap a_0 \bar{b}_i = a_0$ ist, erhält man zufolge (19) durch Summation über die Zellen von \bar{b}_i

$$\mathcal{D} \cap \bar{b}_i = a_0 b'_i.$$

Ist $b_i = b_i^d$, so ist $\bar{b}_i = a_0^{r-1} \bar{b}_i^d$ ein $(r+d)$ -dimensionales abgeschlossenes Simplex, und b_i ist eine $(r+d-1)$ -dimensionale Sphäre, also zu dem durch einen unendlichfernen Punkt abgeschlossenen euklidischen \mathfrak{R}^{r+d-1} homöomorph. $a_0 b'_i$ ist (vgl. R., S. 97, Satz 7) eine $(r+d-2)$ -dimensionale Sphäre, die regulär in b_i liegt und daher b_i nach dem Jordan-Brouwerschen Satz (AH., S. 444) in zwei Gebiete zerlegt:

$$(21) \quad b_i = \mathfrak{M}'_i + a_0 b'_i + \mathfrak{M}''_i;$$

\mathfrak{M}'_i sei der Teil, zu dem a_0^{r-1} gehört. \mathfrak{M}'_i und \mathfrak{M}''_i sind zu einem Simplex a^{r+d-1} homöomorph. Nun ist (R., S. 28, Formel (1))

$$b_i = q'_i b_i + |q'_i|,$$

und durch Kegelbildung mit q'_i als Spitze überträgt sich also die Zerlegung (21) von b_i auf b_i : es ist

$$b_i = q'_i \mathfrak{M}'_i + q'_i (a_0 b'_i) + |q'_i| + q'_i \mathfrak{M}''_i.$$

Wegen $|q'_i| b'_i = \bar{b}'_i$ ist hierin $q'_i (a_0 b'_i) + |q'_i| = a_0 \bar{b}'_i - a_0 b'_i = \mathcal{D}_i$, ferner $q'_i \mathfrak{M}'_i \subset \mathcal{G}'_k$. Wir haben also (20) mit

$$(22) \quad \mathcal{D}'_i = q'_i \mathfrak{M}'_i \subset \mathcal{G}'_k, \quad \mathcal{D}''_i = q'_i \mathfrak{M}''_i,$$

²⁴⁾ Zur Randbildung eines Verbindungsproduktes vgl. R., S. 82.

und \mathfrak{I}_i'' ist homöomorph zu einem α^{r+d} . Ist \mathfrak{d}_i Seite von \mathfrak{d}_i , also $\mathfrak{d}_i \subset \mathfrak{d}_i$, so folgt nach (20), (21), (22)

$$\mathfrak{I}_i'' \subset \mathfrak{M}_i'' \subset \overline{\mathfrak{I}_i''}.$$

Haben also \mathfrak{d}_i und \mathfrak{d}_i eine Seite \mathfrak{d}_i' gemeinsam, so grenzen \mathfrak{I}_i'' und \mathfrak{I}_i'' längs einem $(r+i)$ -dimensionalen Bereich \mathfrak{I}_i'' aneinander:

$$\mathfrak{I}_i'' \subset \overline{\mathfrak{I}_i''} \cap \overline{\mathfrak{I}_i''}.$$

Da \mathfrak{X}_k zusammenhängend ist, folgt hieraus der $(r+1)$ -dimensionale Zusammenhang von $\mathfrak{G}_k = \sum_i \mathfrak{I}_i''$.

Definition. Ist α_0 Nebensimplex eines Komplexes K , so sagen wir, wenn in (13) $n > 1$ ist, daß α_0 seine Umgebung zerlegt, und wenn $n = 1$ ist, daß α_0 seine Umgebung nicht zerlegt.

Beispiel. In einer d -dimensionalen Mannigfaltigkeit zerlegt jedes Simplex α^{d-1} seine Umgebung in zwei Teile; kein α^r mit $r < d-1$ zerlegt seine Umgebung.

[Definition. Ist $r > 1$ und enthält \mathfrak{M} nur Zellen $\alpha^{r'} \in K_1$ mit $r' > r-2$, so erklären wir folgendermaßen einen Streckenkomplex $S = S^r(K_1, \mathfrak{M})$. Zerfällt \mathfrak{M} nach Tilgung aller α^{r-1} in Gebiete \mathfrak{M}_i' , und sind $\alpha_1^{r-1}, \alpha_2^{r-1}, \dots$ diejenigen α^{r-1} aus \mathfrak{M} , die ihre Umgebung zerlegen, so ordnen wir jedem Simplex α^{r-1} einen Punkt x_r und jedem Gebiet \mathfrak{M}_i' einen Punkt y_i zu; ist weiter im Sinne der Formel (13) \mathfrak{G}_i eins der Gebiete aus $\mathfrak{D}_{K_1}(\alpha^{r-1}) - \alpha^{r-1}$ und ist $\mathfrak{G}_i \subset \mathfrak{M}_i'$, so soll dem Gebiet \mathfrak{G}_i eine Strecke $x_r y_i$ entsprechen. Gibt es hiernach mehrere Strecken $x_r y_i$, so wird jede derselben durch einen Teilpunkt z_k in zwei Strecken $x_r z_k + z_k y_i$ zerlegt. S ist der aus den Punkten x_r, y_i, z_k und den Strecken $x_r y_i, x_r z_k, z_k y_i$ bestehende Komplex.

Im Falle (M) ergibt sich durch Spezialisierung die folgende einfachere Formulierung.]

Definition. Ist \mathfrak{P} eine d -dimensionale Mannigfaltigkeit und \mathfrak{M} frei von Zellen bis zur Dimension $d-2$, so werden den $\alpha^{d-1} \subset \mathfrak{M}$ und den $\alpha_i^d \subset \mathfrak{M}$ Punkte x_r, y_i , sowie jeder zwischen α^{d-1} und α_i^d bestehenden Inzidenz eine Strecke $x_r y_i$ zugeordnet. $S^d(K_1, \mathfrak{M})$ sei der von diesen Punkten und Strecken gebildete Komplex.

S wird nicht als in einem Raum eingebettet gedacht. S ist zusammenhängend, weil \mathfrak{M} es ist. Die Zahl der von x_r ausgehenden Strecken ist das n in (13); es ist $n \geq 2$, im Fall (M) ist $n = 2$. [Für \mathfrak{M}_i' ($i = 1, 2, \dots$) gibt es zwei Möglichkeiten:

- \mathfrak{M}_i' ist ein Crundsimplex α^r .
- \mathfrak{M}_i' ist ein $(r+1)$ -dimensional zusammenhängendes²⁵⁾ Gebiet.]

²⁵⁾ Vgl. S. 546.

Der Komplex S wird uns dazu dienen, den Fortsetzungsprozeß der Abbildungsfunktion $g(x)$ durchsichtiger zu machen.

Im folgenden wird mit $\beta^r(X)$ die r -te Bettische Zahl des Komplexes X bezeichnet.

Hilfssatz 6. $S_1 = S^r(K_1, \mathfrak{A})$ existiere und es sei $\beta^1(S_1) > 0$. Dann gibt es eine Unterteilung K_2 von K_1 , die die nicht in \mathfrak{A} liegenden Zellen ungeteilt läßt, und ein Polyeder $\mathfrak{R} \supset \mathfrak{R}'$ mit folgenden Eigenschaften: $\mathfrak{R}' - \mathfrak{R} \subset \mathfrak{A}$, $\mathfrak{R}' - \mathfrak{R}$ besteht aus Zellen von K_2 bis zur Dimension $r-1$ und enthält alle in \mathfrak{A} liegenden Zellen von K_2 bis zur Dimension $r-2$; $\mathfrak{A}' = \mathfrak{A} - (\mathfrak{A} \cap \mathfrak{R}')$ ist zusammenhängend; $S^r(K_2, \mathfrak{A}')$ ist ein Baum.

Beweis. Es sei (unter Unterdrückung der Indizes) xy bzw. xz eine Strecke aus S_1 , die mit dem Koeffizienten 1 in einem Zykel von S_1 auftritt, durch deren Tilgung also $\beta^1(S_1)$ um Eins vermindert würde. Wir werden zunächst K_2 und $\mathfrak{R}' = |N'_2|$ so bestimmen, daß

$$(23) \quad \beta^1(S^r(K_2, \mathfrak{A}')) = \beta^1(S_1) - 1$$

ist und daß $K_2, \mathfrak{A}', N'_2, \mathfrak{R}'$ wieder die Voraussetzungen des Satzes erfüllen. Durch Wiederholung des Verfahrens (Induktion nach $\beta^1(S_1)$) folgt daraus die Behauptung. Wir unterscheiden drei Fälle.

I. Fall (M). Von x gehen nur zwei Strecken xy_1, xy_2 aus. x sei der Zelle a^{r-1} zugeordnet. Wir setzen $K_2 = K_1, \mathfrak{R}' = \mathfrak{R} + \overline{a^{r-1}}$, also $\mathfrak{A}' = \mathfrak{A} - a^{r-1}$.

$S^r(K_2, \mathfrak{A}')$ entsteht aus $S^r(K_1, \mathfrak{A})$ durch Tilgung des Punktes x und der Strecken xy_1, xy_2 ; (23) ist erfüllt.

[II. Fall a) von Seite 557. y entspricht einem Simplex $a^r = u_1 \dots u_{r+1}$. $x = x_1$ entspricht (bei geeigneter Numerierung) der Seite $a^{r-1} = u_2 u_3 \dots u_{r+1}$. Auch gewissen weiteren $(r-1)$ -dimensionalen Seiten

$$a_\lambda^{r-1} = u_1 \dots u_{\lambda-1} u_{\lambda+1} \dots u_{r+1} \\ (\lambda = 1, \dots, l)$$

mögen Punkte $x_\lambda \in S_1$ entsprechen²⁶⁾. Nach der über xy gemachten Voraussetzung ist $l > 1$. Der hier in Betracht kommende Teil von S_1 besteht aus den Punkten y, x_1 und den Strecken $x_\lambda y$ ($\lambda = 1, \dots, l$) (Fig. 2); von den x_λ gehen noch weitere Strecken aus. K_2 entstehe aus K_1 durch Einführung eines neuen Eckpunktes $u_0 \in a^r$ und Zentralunterteilung von a^r . Werden vor-

²⁶⁾ Die übrigen a^{r-1} aus a^r gehören entweder zu \mathfrak{R} oder zerlegen ihre Umgebung nicht; in letzterem Falle ist $\mathfrak{D}_K(a^{r-1}) = a^{r-1} + a^r$.

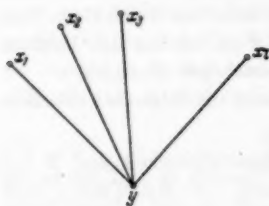


Fig. 2.

läufig von den neu entstehenden Zellen nur diejenigen bis zur Dimension $r-2$ zu \mathcal{N} geschlagen, so ändert sich S folgendermaßen: an die Stelle von y und den Strecken $x_\lambda y$ ($\lambda = 1, \dots, l$) treten y_1, \dots, y_{r+1} (y_β entspricht dem Simplex $a_\beta^{r-1} = u_0 u_1 \dots u_{\beta-1} u_{\beta+1} \dots u_{r+1}$), die Strecken $x_\lambda y_\lambda$ ($\lambda = 1, \dots, l$), weiter die Punkte $x_{\beta\gamma}$ ($0 < \beta < \gamma \leq r+1$; $x_{\beta\gamma}$ entspricht $a_{\beta\gamma}^{r-1} = u_0 u_1 \dots u_{\beta-1} u_{\beta+1} \dots u_{\gamma-1} u_{\gamma+1} \dots u_{r+1}$) und die Strecken $x_{\beta\gamma} y_\beta$, $x_{\beta\gamma} y_\gamma$. Nach I tilgen wir nacheinander alle Punkte $x_{\beta\gamma}$ mit $\beta, \gamma \neq 2$ samt den anstoßenden Strecken; danach hat S im wesentlichen wieder die alten Zusammenhangsverhältnisse (Fig. 3). Werden nun auf dieselbe Weise noch die Strecken $x_{12} y_1$, $x_{12} y_2$ (in Fig. 3 gestrichelt) und der Punkt x_{12} entfernt, so ist die beabsichtigte Trennung zwischen y (jetzt y_2) und x_1 durchgeführt.

III. Fall b) von Seite 557. y entspricht einem $(r+1)$ -dimensional zusammenhängenden Gebiet \mathbb{G} ; x entspreche a_0^{r-1} , gegebenenfalls z einem

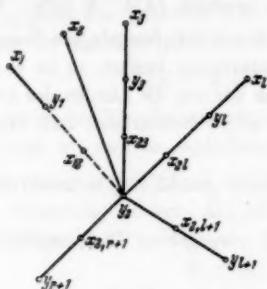


Fig. 3.

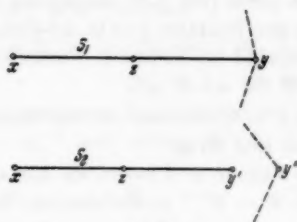


Fig. 4.

Gebiet $\mathbb{G}_k \subset \mathbb{D}_{K_1} (a_0^{r-1}) - a_0^{r-1}$, $\mathbb{G}_k \subset \mathbb{G}$. In y endet noch mindestens eine weitere Strecke (Fig. 4). Hilfssatz 5 liefert eine Unterteilung K_2 von K_1 und ein Polyeder \mathbb{D} aus Zellen von K_2 , das \mathbb{G}_k in zwei Teile zerlegt. \mathcal{N} bestehe aus \mathcal{N} und \mathbb{D} sowie den in \mathbb{G}_k neu entstandenen Zellen bis zur Dimension $r-1$. Letztere Zellen sind Nebensimplexe und zerlegen ihre Umgebung nicht, da sie im Innern eines $(r+1)$ -dimensional zusammenhängenden Gebietes liegen. Da die Begrenzung von \mathbb{G}_k innerhalb \mathbb{G} durch diese Abänderungen nicht berührt wird, ist ebenso wie \mathbb{G}_k auch \mathbb{G} in zwei $(r+1)$ -dimensional zusammenhängende Teilgebiete zerlegt:

$$\mathbb{G} = \mathbb{G}'_k + \mathbb{G} \cap \mathbb{D} + \mathbb{G}''.$$

Wir ordnen den Gebieten $\mathbb{G}'_k, \mathbb{G}''$ Punkte y', y'' in $S_2 = S'(K_2, \mathcal{N})$ zu und haben damit an S die aus Fig. 4 ersichtliche Abänderung vorgenommen; es gilt (23). Damit ist auch III erledigt und der Satz bewiesen.]

Hilfssatz 7. Es sei $r > 1$; \mathbb{E} sei ein euklidischer Raum mit den Koordinaten ξ_1, \dots, ξ_r ; \mathcal{N}^{r-1} sei die Hyperebene $\xi_r = 0$, v ein nicht ver-

schwindender zu \mathcal{R}^{r-1} paralleler Vektor in \mathcal{E}^r , ζ eine positive Zahl, \mathcal{R}^r die Vollkugel $\sum_{i=1}^r \xi_i^2 < \zeta^2$. \mathcal{H}_+ und \mathcal{H}_- seien die durch $\xi_r > 0$ bzw. $\xi_r < 0$ gegebenen Halbräume in \mathcal{E}^r , \mathcal{M} sei die Menge der Punkte mit den Ortsvektoren

$$\mathbf{x} + \lambda \mathbf{v} \quad (x \in \overline{\mathcal{R}^r}, 0 \leq \lambda \leq 1),$$

w sei eine ganz-rationale Zahl.

Dann gibt es eine Abbildung $h(x)$ von $\overline{\mathcal{R}^r}$ in \mathcal{M} mit folgenden Eigenschaften:

$$(24) \quad h(x) = \mathbf{x} + \mathbf{v} \quad \text{für } x \in \mathcal{R}^r,$$

$$(25) \quad h(x) \neq x \quad \text{für } x \in \mathcal{R}^{r-1},$$

$$(26) \quad v(h, \mathcal{H}_- \cap \mathcal{R}^r) = w.$$

Beweis. Da $\overline{\mathcal{R}^r}$ und \mathcal{M} dem euklidischen Raume \mathcal{E}^r angehören, kann man Abbildungen durch Richtungsfelder ersetzen (AH., S. 535). Ist ein Richtungsfeld auf $\overline{\mathcal{R}^r}$ konstruiert, das die durch (24) festgelegten Randwerte und die durch (25), (26) bestimmten Singularitäten besitzt, so ist es leicht, hieraus eine Funktion $h(x)$ zu erhalten; man hat nur die Längen des aus dem Richtungsfeld zu bildenden Vektorfeldes so zu bestimmen, daß (24) und $h(x) \in \mathcal{M}$ für $x \in \mathcal{R}^r$ gilt.

Ist $w = 0$, so setzen wir das Richtungsfeld gemäß (24) konstant über $\overline{\mathcal{R}^r}$ fort und sind fertig.

Im Falle $w \neq 0$ haben wir das auf \mathcal{R}^r vorgegebene Richtungsfeld über $\mathcal{R}^{r-1} = \mathcal{R}^r \cap \mathcal{R}^{r-1}$ so fortzusetzen, daß

$$(26') \quad v_G(h, \mathcal{H}_- \cap \mathcal{R}^r) = w$$

gilt. Dazu bilden wir den Rand \mathcal{A}_- von $\mathcal{A}_- = \mathcal{R}^r \cap \mathcal{H}_-$ durch eine Funktion $\varphi(x)$ so auf eine in \mathcal{E}^r liegende Kugel \mathcal{S}^{r-1} ab, daß für $x \in \mathcal{R}^r \cap \mathcal{H}_-$ $\varphi(x) = y_0 \in \mathcal{S}^{r-1}$ konstant ist und \mathcal{R}^{r-1} durch $\varphi(x)$ topologisch auf $\mathcal{S}^{r-1} - |y_0|$ abgebildet wird. Weiter sei auf \mathcal{S}^{r-1} ein Vektorfeld $v(y)$ erklärt, so daß $v(y_0) = v$ ist und die Kroneckersche Charakteristik dieses Feldes auf \mathcal{S}^{r-1} den Wert $(-1)^r w$ hat. Die Existenz dieses Vektorfeldes bzw. des entsprechenden Richtungsfeldes, also einer Abbildung von \mathcal{S}^{r-1} auf die Richtungskugel vom Grade $(-1)^r w$, folgt aus AH., S. 512, Satz III. Überpflanzen wir das gewonnene Vektorfeld auf \mathcal{A}_- , indem wir dem Punkte $x \in \mathcal{A}_-$ den Vektor $v(\varphi(x))$ zuordnen, und deuten es wieder als Abbildung $h(x)$ um, so sind (24), (25) und (26') erfüllt. Es bleibt noch übrig, das Richtungsfeld über \mathcal{A}_- und $\mathcal{A}_+ = \mathcal{R}^r \cap \mathcal{H}_+$ fortzusetzen; dies ist, da \mathcal{A}_- und \mathcal{A}_+ r -dimensionale Elemente sind, mit je einer Singularität in \mathcal{A}_- und \mathcal{A}_+ nach Hilfssatz 3 möglich. Damit ist Hilfssatz 7 bewiesen.

Haben \mathcal{R}^{r-1} und \mathcal{E}^r die Bedeutung von Hilfssatz 7 und sind w_- und w_+ die zu \mathcal{R}^{r-1} orthogonalen, nach \mathcal{H}_- bzw. \mathcal{H}_+ weisenden Einheitsvektoren

(also $w_+ = \{0, \dots, 0, 1\}$, $w_- = \{0, \dots, 0, -1\}$), so gibt es für $x \in \mathfrak{R}^{r-1}$ eine eindeutig bestimmte Darstellung

$$(27) \quad \eta(x) = x + t(x) + \lambda_-(x)w_- + \lambda_+(x)w_+$$

mit

$$(28) \quad 0 \leq \lambda_- \leq \zeta, \quad 0 \leq \lambda_+ \leq \zeta, \quad \lambda_- \lambda_+ = 0,$$

wo $t(x)$ ein Vektor in \mathfrak{R}^{r-1} ist. t , λ_+ und λ_- hängen stetig von x ab. \mathfrak{C}' ist die Menge der y mit

$$\eta = x + \lambda_- w_- + \lambda_+ w_+,$$

wobei $x \in \mathfrak{R}^{r-1}$ und $\lambda_- \geq 0$, $\lambda_+ \geq 0$, $\lambda_- \lambda_+ = 0$ ist. Faßt man hier \mathfrak{R}^{r-1} als einen Unterraum des euklidischen Raumes \mathfrak{R}^n ($n \geq r$) und w_+ , w_- als irgend zwei verschiedene zu \mathfrak{R}^{r-1} orthogonale Einheitsvektoren in \mathfrak{R}^n auf, so bleibt Hilfssatz 7 gültig. \mathfrak{C}' kann also längs \mathfrak{R}^{r-1} geknickt sein.

Wir werden $h(x)$ im allgemeinen nur für $x \in \mathfrak{R}^{r-1}$ brauchen. Um die Abhängigkeit der Funktion $h(x)$ von den verschiedenen genannten Größen deutlich zu machen, schreiben wir

$$(29) \quad h(x) = h(x, \mathfrak{R}^{r-1}, p, \zeta, v, w_-, w_+, w);$$

dabei sei p der Koordinatenursprung in \mathfrak{R}^{r-1} . In (27) sind $t(x)$, $\lambda_-(x)$ und $\lambda_+(x)$ von w_- und w_+ unabhängig.

Hilfssatz 8. α_0^{r-1} sei ein in \mathfrak{A} enthaltenes Nebensimplex von K_1 , das seine Umgebung zerlegt. $\mathfrak{D}_{K_1}(\alpha_0^{r-1}) - \alpha_0^{r-1}$ und ebenso $\mathfrak{A} - \alpha_0^{r-1}$ zerfalle in n Gebiete ($n > 1$):

$$\begin{aligned} \mathfrak{A} - \alpha_0^{r-1} &= \mathfrak{A}_1 + \dots + \mathfrak{A}_n, \\ \mathfrak{D}_{K_1}(\alpha_0^{r-1}) - \alpha_0^{r-1} &= \mathfrak{G}_1 + \dots + \mathfrak{G}_n, \\ \mathfrak{G}_r &\subset \mathfrak{A}_r, \end{aligned} \quad (r = 1, \dots, n).$$

\mathfrak{R}^{r-1} sei der von α_0^{r-1} aufgespannte Raum. w_r ($r = 1, \dots, n$) sei ein zu \mathfrak{R}^{r-1} orthogonaler Einheitsvektor, so daß für $x \in \alpha_0^{r-1}$ und hinreichend kleines $\lambda > 0$ die durch

$$\eta = x + \lambda w_r$$

bestimmten Punkte y in \mathfrak{G}_r liegen. $g^*(x)$ sei erklärt als stetige Fortsetzung (nach Hilfssatz 3) von $g(x)$ über $\mathfrak{R} + \alpha_0^{r-1}$. \mathfrak{D} sei ein Gebiet in α_0^{r-1} , v ein Vektor in \mathfrak{R}^{r-1} ; für $x \in \mathfrak{D}$ sei $g^*(x) = x + v$, $g^*(x) \in \alpha_0^{r-1}$. w sei eine ganzrationale Zahl. Dann lassen sich $p \in \mathfrak{D}$ und $\zeta > 0$ so bestimmen, daß die Funktion

$$(30) \quad g_2(x) = \begin{cases} h(x, \mathfrak{R}^{r-1}, p, \zeta, v, w_1, w_2, w) & \text{für } x \in \alpha_0^{r-1}, \varrho(p, x) < \zeta, \\ g^*(x) & \text{sonst für } x \in \mathfrak{R} + \alpha_0^{r-1} \end{cases}$$

die Bedingung

$$(31) \quad g_2(x) \in \mathfrak{B}_K(x)$$

erfüllt; wird

$$(32) \quad v_G(g^*, \mathfrak{A}_v) = w_v, \quad (v = 1, \dots, n)$$

gesetzt, so ist

$$(33) \quad v_G(g_2, \mathfrak{A}_1) = w_1 + w$$

$$(34) \quad v_G(g_2, \mathfrak{A}_2) = w_2 - w$$

[sowie im Falle $n > 2$

$$(35) \quad v_G(g_2, \mathfrak{A}_3) = w_3].$$

Beweis. I. Wir bemerken, daß die Vektoren w , mit den geforderten Eigenschaften sich stets angeben lassen; man hat nur ein $y \in \mathfrak{G}$, zu wählen und w , als den von \mathfrak{R}^{r-1} aus in den durch \mathfrak{R}^{r-1} und y , bestimmten Halbraum weisenden Einheitsvektor zu bestimmen. $p \in \mathfrak{D}$ werde beliebig gewählt; dann läßt sich ζ stets so bestimmen, daß die durch $\varrho(p, x) < \zeta$, $x \in \mathfrak{R}^{r-1}$ gegebene $(r-1)$ -dimensionale Kreisscheibe \mathfrak{R}^{r-1} in \mathfrak{D} enthalten ist und daß für $x \in \mathfrak{R}^{r-1}$, $0 \leq \lambda \leq 1$ und $0 < \lambda_r \leq \zeta$ die Punkte y mit

$$y = x + \lambda v + \lambda_r w, \quad (v = 1, \dots, n)$$

dem Gebiet \mathfrak{G} , angehören. Dann ist die Funktion (30) sinnvoll und stetig und erfüllt (31).

[Um zunächst (35) zu beweisen, bilden wir eine Hilfsfunktion $g_1(x)$ und zeigen

$$(36) \quad v_G(g_1, \mathfrak{A}_3) = v_G(g_2, \mathfrak{A}_3),$$

$$(37) \quad v_G(g_1, \mathfrak{A}_3) = w_3.$$

II. Es werde $g_1(x)$ dadurch erklärt, daß in der Definition (30) von $g_2(x)$ rechts w_2 durch w_1 ersetzt wird. Zwar war in (29) für die Anwendung von

Hilfssatz 7 $w_- \neq w_+$ vorausgesetzt, doch sieht man an dem Ausdruck (27), daß

$h(x)$ für $w_- = w_+ = w_1$ sinnvoll bleibt.

Wegen (28) ist $0 \leq \lambda_- + \lambda_+ \leq \zeta$ und daher $g_1(x) \in \mathfrak{a}_0^{r-1} + \mathfrak{G}_1$ für $x \in \mathfrak{R}^{r-1}$. \mathfrak{F} sei die Menge der $y \in \mathfrak{G}_3$ mit $\varrho(p, y) < \zeta$. Jedes solche y ist eindeutig darzustellen durch

$$(38) \quad y = \tilde{y} + \sigma \times w,$$

wenn $\tilde{y} \in \mathfrak{R}^{r-1}$, w irgendein zu \mathfrak{R}^{r-1} orthogonaler ins Innere von \mathfrak{G}_3 weisender Einheitsvektor, $0 < \sigma < 1$ und $\sigma = \sqrt{\zeta^2 - (\varrho(p, \tilde{y}))^2}$ ist. \tilde{y} , σ und w hängen stetig von y ab und durchlaufen unabhängig voneinander alle Werte



Fig. 4.

der angegebenen Bereiche. Wir setzen nun g_1 und g_2 in bestimmter Weise über $\bar{\mathfrak{F}}$ fort. Die Begrenzung von $\bar{\mathfrak{F}}$ in \mathfrak{P} ist

$$(39) \quad \bar{\mathfrak{F}} = \bar{\mathfrak{R}^{r-1}} + \mathfrak{L},$$

dabei ist \mathfrak{L} die Menge der y mit

$$(40) \quad \eta = \bar{\eta} + \kappa w,$$

wo $\bar{\eta}$ und w dieselben Wertebereiche wie in (38) durchlaufen. Es ist nach (27), (30) und der Definition von $g_i(x)$ für $x \in \bar{\mathfrak{R}^{r-1}}$

$$(41) \quad g_i(x) = x + t(x) + \lambda_-(x) w_1 + \lambda_+(x) w_i \quad (i = 1, 2);$$

wir setzen

$$(42) \quad g_i(y) = \bar{\eta} + \sigma v + (1 - \sigma)(t(\bar{y}) + \lambda_-(\bar{y}) w_1 + \lambda_+(\bar{y}) w_i) \quad \text{für } y \in \bar{\mathfrak{F}},$$

$$(43) \quad g_i(y) = \bar{\eta} + v \quad \text{für } y \in \mathfrak{L} \quad (i = 1, 2).$$

$g_i(y)$ ist nach (42) stetig auf $\bar{\mathfrak{F}}$ und nach (43) auf \mathfrak{L} ; geht man in (42) zur Grenze über ($\sigma \rightarrow 0$ bzw. $\sigma \rightarrow 1$ bzw. $\bar{y} \rightarrow \bar{\mathfrak{R}^{r-1}}$), so erkennt man durch Vergleich mit (41) und (43) die Stetigkeit auf $\bar{\mathfrak{F}}$. Es ist $g_i(y) \in \mathfrak{D}_{K_1}(a_0^{r-1})$, daher

$$(44) \quad g_i(y) \in \mathfrak{B}_K(y).$$

Auf $\bar{\mathfrak{F}}$ sind g_1 und g_2 fixpunktfrei; für $y \in \bar{\mathfrak{R}^{r-1}}$ folgt dies aus der Definition von $g_i(y)$ und aus Hilfssatz 7, für $y \in \mathfrak{L} + \bar{\mathfrak{F}}$ folgt es aus (43) und (42), denn nach diesen Formeln ist stets $g_i(y) \in \mathfrak{R}^{r-1} + \mathfrak{G}_1 + \mathfrak{G}_2$, also nie $g_i(y) \in \mathfrak{G}_3$ ($i = 1, 2$). Man hat somit

$$(45) \quad v(g_1, \bar{\mathfrak{F}}) = v(g_2, \bar{\mathfrak{F}}) = 0.$$

Weiter ist die Begrenzung von $\mathfrak{A}_3 - \bar{\mathfrak{F}}$ in \mathfrak{P} zufolge (39)

$$(\mathfrak{A}_3 - \bar{\mathfrak{F}}) = \mathfrak{A}_3 - \mathfrak{R}^{r-1} + \mathfrak{L},$$

es ist also

$$(46) \quad g_1(y) = g_2(y) \quad \text{für } y \in (\mathfrak{A}_3 - \bar{\mathfrak{F}}).$$

Aus (44) und (46) folgt nach Hilfssatz 4²⁷⁾

$$(47) \quad v_G(g_1, \mathfrak{A}_3 - \bar{\mathfrak{F}}) = v_G(g_2, \mathfrak{A}_3 - \bar{\mathfrak{F}});$$

aus (45), worin v durch v_G ersetzt werden darf, und (47) folgt (36).

III. Um (37) zu beweisen, deformieren wir unter Vermeidung von Fixpunkten g_1 stetig in g^* dort, wo beide Abbildungen voneinander verschieden sind, also auf \mathfrak{R}^{r-1} . Es ist dort

$$g^*(x) = x + v,$$

$$g_1(x) = x + t(x) + (\lambda_-(x) + \lambda_+(x)) w_1;$$

²⁷⁾ Hilfssatz 4 ist freilich erst anwendbar, nachdem die durch \mathfrak{L} bewirkte Zerlegung von \mathfrak{A}_3 zu einer krummen Simplicialzerlegung ausgebaut ist, die eine Unterteilung von K ist. Wir übergehen den leichten Beweis, daß dies stets möglich ist.

erklärt man $g(x; \sigma)$ für $\sigma = 0, 1, 2, 3$ durch

$$\begin{aligned} g(x; 0) &= g_1(x); & g(x; 1) &= x + t(x) + \zeta w_1; \\ g(x; 2) &= x + v + \zeta w_1; & g(x; 3) &= g^*(x) \end{aligned}$$

und für die Zwischenwerte durch lineare Interpolation, so ist für $0 \leq \sigma \leq 3$, $x \in \mathcal{R}^{r-1}$ $g(x; \sigma)$ stetig in x und σ , $g(x; \sigma) \in \mathfrak{B}_K(x)$ und $g(x; \sigma) \neq x$. Daraus folgt nach S. 554 die Identität

$$v_G(g_1, \mathfrak{A}_3) = v_G(g^*, \mathfrak{A}_3),$$

also (37). Aus (36) und (37) folgt die Gleichung (35), die damit bewiesen ist. — Zum Nachweis von (33) führen wir eine weitere Hilfsfunktion $g_{n+1}(x)$ ein und zeigen

$$(48) \quad v(g_{n+1}, \mathfrak{A}_1) = w_1 + w,$$

$$(49) \quad v_G(g_{n+1}, \mathfrak{A}_1) = v_G(g_2, \mathfrak{A}_1).$$

IV. w_{n+1} sei ein zu \mathcal{R}^{r-1} orthogonaler Einheitsvektor, so daß die durch

$$(50) \quad \eta = \bar{\eta} + \lambda w_{n+1} \quad (\bar{\eta} \in \mathfrak{a}_0^{r-1}, \lambda > 0)$$

bezeichneten Punkte nicht zu \mathfrak{P} gehören²⁸⁾, und α' sei ein in dem durch \mathcal{R}^{r-1} und w_{n+1} bestimmten Halbraume liegendes von \mathfrak{a}_0^{r-1} berandetes Simplex²⁹⁾. α' sei so gewählt, daß jeder Punkt $y \in \mathfrak{A}_{n+1} = \alpha' - \mathfrak{a}_0^{r-1}$ die Gestalt (50) hat, und daß durch (30) eine Abbildungsfunktion g_{n+1} in $\mathfrak{P} + \alpha'$ erklärt wird, wenn man dort w_2 durch w_{n+1} ersetzt. Durch

$$g^*(y) = g^*(\bar{y}) \quad \text{für } y \in \mathfrak{A}_{n+1}$$

setzen wir g^* und analog auch g_1 und g_2 über \mathfrak{A}_{n+1} stetig und fixpunktfrei fort, haben also

$$(51) \quad v(g_1, \mathfrak{A}_{n+1}) = v(g_2, \mathfrak{A}_{n+1}) = v(g^*, \mathfrak{A}_{n+1}) = 0.$$

Zwecks Fortsetzung von g_{n+1} über \mathfrak{A}_{n+1} gehen wir auf die ursprünglich in Hilfssatz 7 konstruierte Funktion $h(x)$ zurück, die auf einer Vollkugel $\mathcal{R}^r \subset \mathbb{E}^r$ definiert war; indem wir das Vektorfeld aus der dortigen Halbkugel $\mathfrak{A}_+ = \mathcal{R}^r \cap \mathfrak{H}_+$ sinngemäß überpflanzen, gewinnen wir die Funktionswerte $g_{n+1}(y)$ für alle y aus der r -dimensionalen Halbkugel \mathfrak{H} , die in dem von \mathcal{R}^{r-1} berandeten und α' enthaltenden Halbraume über der Kreisscheibe \mathcal{R}^{r-1} errichtet ist. Nach Hilfssatz 7 ist

$$(52) \quad v(g_{n+1}, \mathfrak{H}) = -w.$$

²⁸⁾ Nötigenfalls muß man hier die Dimension des einbettenden Raumes um Eins erhöhen.

²⁹⁾ Die Hinzunahme von α' zu \mathfrak{P} ist zur Anwendung von Hilfssatz 7 erforderlich, da sich unter den \mathfrak{G}_r nicht notwendig ein r -dimensionales Gebiet befindet.

Um g_{n+1} über \bar{a}^r fortzusetzen, verfahren wir ähnlich wie oben bei g^* , setzen jedoch für $y \in \bar{a}^r - \bar{a}^{r-1} - \mathfrak{F}$

$$g_{n+1}(y) = g_{n+1}(\bar{y}) + \kappa w_{n+1},$$

wobei $\kappa = 0$ für $\bar{y} \notin \mathcal{R}^{r-1}$ und $\kappa = \sqrt{\zeta^2 - (\varrho(p, \bar{y}))^2}$ für $\bar{y} \in \mathcal{R}^{r-1}$ zu setzen ist. g_{n+1} ist in $\mathfrak{A}_{n+1} - \mathfrak{F}$ fixpunktfrei, also ist nach (52)

$$(53) \quad v(g_{n+1}, \mathfrak{A}_{n+1}) = -w.$$

Da mit (35) aus Symmetriegründen auch

$$(54) \quad v_G(g_{n+1}, \mathfrak{A}_r) = w_r \quad (2 \leq r \leq n)$$

bewiesen ist und bei Fortsetzung von g^* und g_{n+1} nach Hilfssatz 3 über \mathfrak{P}

$$(55) \quad \sum_{r=1}^{n+1} v(g_{n+1}, \mathfrak{A}_r) = \sum_{r=1}^{n+1} v(g^*, \mathfrak{A}_r) = v_G(g^*, \mathfrak{P} + \bar{a}^r) = \chi(\mathfrak{P} + \bar{a}^r)$$

ist, folgt nach (32), (51), (53)

$$v(g_{n+1}, \mathfrak{A}_1) + \sum_{r=2}^n w_r - w = \sum_{r=1}^n w_r,$$

also (48).

V. Zum Beweise von (49) erklären wir wie oben unter II \bar{y} , σ , κ , w , \mathfrak{F} und \mathfrak{Q} für alle $y \in \mathfrak{G}_1$ mit $\varrho(p, y) < \zeta$, indem wir das dortige \mathfrak{G}_3 durch \mathfrak{G}_1 ersetzen (vgl. Fig. 5, in der ebenfalls \mathfrak{G}_1 statt \mathfrak{G}_3 zu setzen ist). Entsprechend (41), (42), (43) gewinnen wir die Formeln

$$(56) \quad g_i(x) = x + t(x) + \lambda(x)w_1 + \lambda_+(x)w_i \quad (i = 2, n+1; x \in \overline{\mathcal{R}^{r-1}}),$$

$$(57) \quad g_i(y) = \bar{y} + \sigma v + (1 - \sigma)(t(\bar{y}) + \lambda_-(\bar{y})w_1 + \lambda_+(\bar{y})w_i) \\ (i = 2, n+1; y \in \mathfrak{F}),$$

$$(58) \quad g_i(y) = \bar{y} + v \quad (i = 2, \dots, n+1; y \in \mathfrak{Q}).$$

Es sei y_0 Fixpunkt bei g_2 oder g_{n+1} , also $y_0 \in \mathfrak{F}$. Dann muß in einer Umgebung von y_0

$$(59) \quad \lambda_+(\bar{y}) = 0$$

sein, da sonst nach (57) $g_i(y_0) = y_0$ in $\overline{\mathfrak{G}_2}$ bzw. \bar{a}^r liegen müßte. Aus (59) und (57) folgt, daß g_2 und g_{n+1} in \mathfrak{F} nur gemeinsame Fixpunkte haben und daß in einer Umgebung eines solchen stets $g_2(y) = g_{n+1}(y)$ ist. Daraus schließen wir

$$v(g_2, \mathfrak{F}) = v(g_{n+1}, \mathfrak{F}).$$

Da man analog zu (47)

$$v_G(g_2, \mathfrak{A}_1 - \mathfrak{F}) = v_G(g_{n+1}, \mathfrak{A}_1 - \mathfrak{F})$$

hat, gilt in der Tat (49). Aus (48) und (49) folgt die Behauptung (33).

VI. Da nach (35) aus Symmetriegründen

$$v_G(g_2, \mathfrak{A}_r) = w_r \quad (3 \leq r \leq n)$$

ist, findet man mit Hilfe von (32), (48) und (33) durch Summation analog (55), jedoch unter Fortlassung von \mathfrak{A}_{n+1} , die letzte noch ausstehende Behauptung (34). Damit ist Hilfssatz 8 vollständig bewiesen; wir fügen jedoch noch einen verkürzten Beweis für Fall (M) an.]

VII. Im Falle (M) ist $n = 2$, $\mathfrak{D}_{K_1}(a_0^{r-1}) = a_0^{r-1} + a_1^r + a_2^r$, $a_1^r \subset \mathfrak{A}_1$, $a_2^r \subset \mathfrak{A}_2$. w_1 und w_2 sind eindeutig bestimmt und legen in Verbindung mit \mathfrak{R}^{r-1} einen Raum \mathfrak{E}^r fest, in dem $\mathfrak{D}_{K_1}(a_0^{r-1})$ ein Gebiet ist. Die Menge \mathfrak{R}^r der $y \in \mathfrak{P}$ mit $\varrho(p, y) < \zeta$ ist zu einer Kugel homöomorph und sogar isometrisch; es sei $\mathfrak{R}^r \cap a_2^r = \mathfrak{F}_+$, $\mathfrak{R}^r \cap a_2^r = \mathfrak{L}$. $g^*(x)$ werde nach Hilfssatz 3 über \mathfrak{A} fortgesetzt. Durch eine Funktion $\varphi(x)$ werde \bar{a}_2^r auf $\bar{a}_2^r - \mathfrak{R}^{r-1} - \mathfrak{F}_+$ und gleichzeitig \mathfrak{R}^{r-1} auf \mathfrak{L} topologisch abgebildet, so daß die übrigen Randpunkte von a_2^r fest bleiben. φ kann so gewählt werden, daß $\varphi(x) \neq x$ nur in einer gewissen Umgebung $\mathfrak{U} \supset \mathfrak{F}_+$ gilt, wo $g^*(x) \in \mathfrak{D}_{K_1}(a_0^{r-1}) \subset \mathfrak{E}^r$ ist und man daher g^* ein Vektorfeld in \mathfrak{E}^r zuordnen kann. Indem wir dieses Vektorfeld bei der Abbildung $x \rightarrow \varphi(x)$ durch Parallelverschiebung überpflanzen, entsteht auf $\bar{a}_2^r - \mathfrak{R}^{r-1} - \mathfrak{F}_+$ eine Abbildungsfunktion $g^{**}(x)$ mit

$$(60) \quad v(g^{**}, a_2^r - \mathfrak{F}_+) = v(g^*, a_2^r).$$

Setzt man

$$(61) \quad g_2(x) = \begin{cases} g^{**}(x) & \text{für } x \in \bar{a}_2^r - \mathfrak{F}_+, \\ g^*(x) & \text{für } x \in \mathfrak{A}_2 - a_2^r \end{cases}$$

und erklärt $g_2(x)$ auf \mathfrak{F}_+ durch die in Hilfssatz 7 konstruierte Funktion $h(x)$, so besteht in \mathfrak{L} ein stetiger Übergang, und $g_2(x)$ hat auf a_0^{r-1} die vorgeschriebenen Werte; man hat nach (60) und (61)

$$v(g_2, \mathfrak{A}_2 - \mathfrak{F}_+) = v(g^*, \mathfrak{A}_2 - a_2^r) + v(g^{**}, a_2^r - \mathfrak{F}_+) = w_2,$$

ferner nach Hilfssatz 7

$$v(g_2, \mathfrak{F}_+) = -w,$$

daher gilt (34), und wegen $v_G(g_2, \mathfrak{A}_1 + \mathfrak{A}_2) = v_G(g^*, \mathfrak{A}_1 + \mathfrak{A}_2)$ auch (33).

Hilfssatz 9. $S^r(K_1, \mathfrak{A})$ sei ein Baum. Es sei $a_0^{r-1} \in K_1$, $a_0^{r-1} \subset \mathfrak{A}$; $\mathfrak{A} - a_0^{r-1}$ zerfalle in n Gebiete $\mathfrak{A}_1, \dots, \mathfrak{A}_n$ ($n > 1$). Dann läßt sich $g(x)$ so über a_0^{r-1} mit (3) und (4) fortsetzen, daß $v_G(g, \mathfrak{A}_2) = \dots = v_G(g, \mathfrak{A}_n) = 0$ wird.

Beweis. Da $S^r(K_1, \mathfrak{A})$ ein Baum ist, ist a_0^{r-1} Nebencomplex und $\mathfrak{D}_{K_1}(a_0^{r-1}) - a_0^{r-1}$ zerfällt in dieselbe Anzahl von Gebieten wie $\mathfrak{A} - a_0^{r-1}$. Die auf \mathfrak{P} , \mathfrak{A} und a_0^{r-1} bezüglichen Voraussetzungen von Hilfssatz 8 sind also erfüllt. $g_a(x)$ sei eine auf $\mathfrak{A} + a_0^{r-1}$ erklärte Funktion, die Fortsetzung von $g(x)$ nach Hilfssatz 3 ist, und g_a sei so gewählt, daß es einen Punkt $p \in a_0^{r-1}$ mit

$$(62) \quad g_a(p) \in a_0^{r-1}$$

gibt, was offenbar leicht zu erreichen ist. Weiter wählen wir Simplexe a_1^{r-1} , a_2^{r-1} mit $p \in a_2^{r-1} \subset a_1^{r-1} \subset a_0^{r-1}$, so daß a_2^{r-1} , a_1^{r-1} , a_0^{r-1} paarweise punktfremd sind und

$$g_a(x) \in \mathfrak{D}_{K_1}(p) \quad \text{für } x \in \overline{a_1^{r-1}}$$

ist. Es kann hiernach g_a auf a_1^{r-1} in bekannter Weise durch ein Richtungsfeld ersetzt werden. $\overline{a_1^{r-1}} - |p|$ werde nun topologisch auf $\overline{a_1^{r-1}} - \overline{a_2^{r-1}}$ abgebildet, so daß $\overline{a_1^{r-1}}$ punktweise fest bleibt, und dabei das Richtungsfeld überpflanzt; es entsteht ein neues Feld, das auf $a_1^{r-1} - \overline{a_2^{r-1}}$ erklärt ist und nach stetiger Ergänzung auf a_2^{r-1} konstant ist. Wir setzen es konstant über $\overline{a_2^{r-1}}$ fort und bilden daraus wieder eine Abbildung, die wir mit $g^*(x)$ bezeichnen, und die stetige Fortsetzung von g ist; dabei können nach (62) die Verschiebungsvektoren auf a_2^{r-1} konstant, etwa gleich v , und zwar so gewählt werden, daß die Bildpunkte in a_0^{r-1} liegen:

$$g^*(x) = x + v, \quad g^*(x) \in a_0^{r-1} \quad \text{für } x \in a_2^{r-1}.$$

Danach erfüllen g , g^* und $\mathfrak{D} = a_2^{r-1}$ die Voraussetzungen von Hilfssatz 8; die Zahlen w_ν ($\nu = 1, \dots, n$) mögen die Bedeutung (32) haben. $\mathfrak{D}_2, \dots, \mathfrak{D}_n$ seien $n-1$ paarweise punktfremde Teilgebiete von a_2^{r-1} . Wir wenden Hilfssatz 8 an mit \mathfrak{D}_2 statt \mathfrak{D} , $w = w_2$ und gewinnen eine Funktion $g_2(x)$ mit

$$v_G(g_2, \mathfrak{A}_2) = 0, \quad v_G(g_2, \mathfrak{A}_\nu) = w_\nu \quad (\nu = 3, \dots, n).$$

Im Falle (M) ist wegen $n = 2$ hier der Beweis zu Ende.

[Erneute Anwendung des gleichen Satzes mit $w = w_3$, \mathfrak{A}_3 statt \mathfrak{A}_2 , g_3 statt g_2 , \mathfrak{D}_3 statt \mathfrak{D} , g_1 statt g^* ergibt g_3 mit

$$v_G(g_3, \mathfrak{A}_2) = v_G(g_3, \mathfrak{A}_3) = 0, \quad v_G(g_3, \mathfrak{A}_\nu) = w_\nu \quad (\nu = 4, \dots, n).$$

Durch Wiederholung des Verfahrens (Induktion nach n) gelangt man schließlich zu einer Funktion g_n mit

$$v_G(g_n, \mathfrak{A}_\nu) = 0 \quad (\nu = 2, \dots, n);$$

g_n ist die gewünschte Fortsetzung von g , und wir schreiben $g = g_n$. Damit ist der Satz bewiesen.]

Hilfssatz 10. Es sei $r > 1$; \mathfrak{A} enthalte ein r -dimensionales Simplex, aber keins von geringerer Dimension als $r-1$. Dann gibt es eine Unterteilung K_2 von K_1 , die die nicht zu \mathfrak{A} gehörenden Simplexe ungeteilt läßt, und man kann $g(x)$ mit (3) und (4) über die in \mathfrak{A} enthaltenen Zellen von K_2 bis zur Dimension $r-1$ derart fortsetzen, daß der Rest von \mathfrak{A} in Gebiete $\mathfrak{A}_1, \dots, \mathfrak{A}_n$ mit $v_G(g, \mathfrak{A}_\nu) = 0$ ($\nu = 2, \dots, n$) zerfällt.

Beweis³⁰⁾. Alle $(r-1)$ -dimensionalen Simplexe in \mathfrak{A} sind Nebensimplexe. [Über diejenigen von ihnen, die ihre Umgebung nicht zerlegen, setzen wir alsbald $g(x)$ nach Hilfssatz 3 fort.] Wir bilden $S^r(K_1, \mathfrak{A})$ und wenden Hilfssatz 6 an; es wird also K_1 durch eine Unterteilung K_2 ersetzt und \mathfrak{A} durch ein kleineres Gebiet \mathfrak{A}' , so daß $S^r(K_2, \mathfrak{A}')$ ein Baum ist. Jetzt ist Hilfssatz 9 (mit \mathfrak{A}' statt \mathfrak{A} , K_2 statt K_1) anwendbar; er liefert eine Fortsetzung von g über ein Simplex $\alpha_0^{r-1} \subset \mathfrak{A}'$, so daß von den neu entstehenden Gebieten \mathfrak{A}_v nur ein einziges, das willkürlich gewählt werden kann, eine von Null verschiedene Grenzf়ixpunktzahl erhält. Die dabei mit $S^r(K_2, \mathfrak{A}')$ erfolgte Abänderung besteht darin, daß ein Punkt x_i (der der Zelle α_0^{r-1} zugeordnet war) samt den anstoßenden n_1 Strecken $x_i y_i$ bzw. Streckenzügen $x_i z_k + |z_k| + z_k y_i$ getilgt wurde, und daß $S^r(K_2, \mathfrak{A}')$ nun in n_1 Bäume $S^r(K_2, \mathfrak{A}_v)$ ($v = 1, \dots, n_1$) zerfallen ist. Es ist natürlich $v_G(g, \mathfrak{A}) = v_G(g, \mathfrak{A}')$. Durch erneute Anwendung von Hilfssatz 9 auf die $S^r(K_2, \mathfrak{A}_v)$, soweit sie mehr als einen Punkt enthalten, gelangt man durch Induktion zu n Komplexen, die nur aus je einem Punkt bestehen. Dies sind die y -Punkte von $S^r(K_2, \mathfrak{A}')$, sie entsprechen also gewissen Gebieten, die von Zellen bis zur Dimension $r-1$ frei sind. Eins dieser Gebiete hat die Grenzf়ixpunktzahl von \mathfrak{A}' , für die übrigen ist $v_G = 0$.

Satz 1. *Haben \mathfrak{A} und $g(x)$ die auf S. 553 angegebene Bedeutung, und ist \mathfrak{A} zweidimensional zusammenhängend, so läßt sich $g(x)$ mit (3) so über \mathfrak{A} fortsetzen, daß höchstens ein Fixpunkt entsteht.*

Beweis. Man setze g über die nulldimensionalen Zellen von \mathfrak{A} nach Hilfssatz 3 fort und wende auf das Restgebiet Hilfssatz 10 mit $r = 2$ an. Die neu entstehenden Gebiete sind (vgl. S. 557) zum Teil Dreiecke, zum Teil dreidimensional zusammenhängende Gebiete. Über jene wird g nach Hilfssatz 3 fortgesetzt, auf diese wird Hilfssatz 10 mit $r = 3$ angewendet usw.³¹⁾ So werden nacheinander die 2-, 3-, ...-dimensionalen Grundsimplxe isoliert und $g(x)$ darauf definiert. Ein einziges von ihnen erhält nach Konstruktion eine Fixpunktzahl $v(g, \alpha') = v_G(g, \alpha') = v_G(g, \mathfrak{A})$, die von Null verschieden sein kann, und daher eventuell einen Fixpunkt.

Satz 1 wird für die Anwendung in § 2 wichtig sein. — Es ist leicht zu erkennen, daß die Lage des Fixpunktes, der im Falle $v_G(g, \mathfrak{A}) \neq 0$ entsteht, weitgehend willkürlich vorgeschrieben werden kann. Bei jeder Zerfällung eines Gebietes \mathfrak{A} nach Hilfssatz 9 geht ja die Fixpunktzahl von \mathfrak{A} auf ein willkürlich mit \mathfrak{A}_1 bezeichnetes Teilgebiet über, und ist schließlich $v_G(g, \alpha') \neq 0$, so kann innerhalb α' der Fixpunkt noch willkürlich gewählt werden. Berücksichtigt

³⁰⁾ Im Falle (M) ist nur dann etwas zu beweisen, wenn r die Dimension von \mathfrak{P} ist. Es kann hier $K_2 = K_1$ gewählt werden. Die Gebiete $\mathfrak{A}_1, \dots, \mathfrak{A}_n$ sind genau die nicht in \mathfrak{A} enthaltenen Grundsimplxe von K_1 .

³¹⁾ Im Falle (M) ist nur eine einmalige Anwendung von Hilfssatz 10 erforderlich.

man, daß bei geeigneter Zerlegung von \mathfrak{P} bzw. \mathfrak{A} jeder reguläre Punkt³²⁾ innerer Punkt eines Grundsimplex werden kann, so erkennt man: *im Falle $v_0(g, \mathfrak{A}) \neq 0$ kann der in \mathfrak{A} entstehende Fixpunkt als regulärer Punkt von \mathfrak{A} willkürlich vorgegeben werden.*

Durch Spezialisierung auf den Fall, daß \mathfrak{R} leer ist, gewinnt man aus Satz 1 den folgenden Hauptsatz.

Satz 2. *Ist \mathfrak{P} ein zweidimensional zusammenhängendes Polyeder, so gibt es eine Deformation von \mathfrak{P} in sich mit nur einem Fixpunkt, dieser hat die Vielfachheit $\chi(\mathfrak{P})$. Ist $\chi(\mathfrak{P}) = 0$, so gibt es auch eine fixpunktfreie Deformation.*

Im Sinne der in der Einleitung benutzten Bezeichnungen läßt sich Satz 2 auch so aussprechen:

Für die Abbildungsklasse der Identität bei zweidimensional zusammenhängenden Polyedern gilt $m = \mu$, d. h. die geometrische Fixpunktmindestzahl ist gleich der Zahl der wesentlichen Fixpunktklassen.

§ 2.

Beliebige Abbildungsklasse.

Für diesen Paragraphen wird über das Polyeder \mathfrak{P} die nachstehende Voraussetzung gemacht:

D. $\mathfrak{P} = |K|$ sei dreidimensional zusammenhängend. Jeder Weg in \mathfrak{P} , der zwei reguläre Punkte verbindet, sei homotop³³⁾ zu einem solchen, der ganz in einem dreidimensional zusammenhängenden Gebiet in \mathfrak{P} verläuft.

Die Bedingung **D** fordert mehr als den dreidimensionalen Zusammenhang. Wir können sie für die zu beweisenden Sätze durch eine noch schärfere, aber bequemere Bedingung **D'** ersetzen.

Da man ein mindestens d -dimensionales Polyeder d -dimensional zusammenhängend nennt, wenn es nicht durch ein höchstens $(d-2)$ -dimensionales Teilpolyeder zerlegt werden kann, wird durch die auf S. 557 eingeführte Ausdrucksweise, daß ein Nebensimplex seine Umgebung zerlegt, die folgende Verschärfung jener Definition nahegelegt: *Ein mindestens d -dimensionales Polyeder $\mathfrak{P} = |K|$ heißt lokal d -dimensional zusammenhängend, wenn K kein höchstens $(d-2)$ -dimensionales Nebensimplex enthält, das seine Umgebung zerlegt.* Gleichwertig ist die Formulierung: *\mathfrak{P} heißt lokal d -dimensional zusammenhängend, wenn jedes Gebiet in \mathfrak{P} d -dimensional zusammenhängend ist.* Für Mannigfaltigkeiten sowie für $d = 1$ sind d -dimensionaler und lokaler.

³²⁾ Vgl. W., S. 663.

³³⁾ Vgl. W., S. 665.

d -dimensionaler Zusammenhang gleichbedeutend; im übrigen besagt D weniger als die folgende Bedingung:

D' . \mathfrak{P} sei ein lokal dreidimensional zusammenhängendes Polyeder.

Benutzt wird, wie gesagt, nur die Bedingung D .

Wir nehmen weiter eine Abbildungsfunktion $f(x)$ als gegeben an und setzen von hier ab bis Hilfssatz 15 einschließlich voraus: $f(x)$, abgekürzt mit x' bezeichnet, sei eine stetige Selbstabbildung von \mathfrak{P} mit nur regulären Fixpunkten, die in Grundsimplexen von K liegen. p und q seien zwei Fixpunkte bei $f(x)$, die zur gleichen Fixpunktklasse gehören.

Hilfssatz 11. Es gibt eine Unterteilung K_1 von K , Simplexe $a_0^{r_0}, a_1^{r_1}, a_r^{d_r}$ aus K_1 und Punkte q_r ($r = 1, \dots, k$) mit folgenden Eigenschaften: $q_r \in a_r^{d_r}$, $p \in a_0^{r_0}$, $q \in a_k^{r_k}$, $\overline{a_r^{d_r}} = \overline{a_{r-1}^{r_{r-1}}} \cap \overline{a_r^{r_r}}$, $q, q_{r+1} \subset a_r^{r_r}$, $d_r \geq 2$; die $a_r^{r_r}$ sind Grundsimplxe, alle $a_r^{r_r}$ und $a_r^{d_r}$ sind paarweise verschieden; das Gebiet $\mathfrak{G} = \sum_{r=1}^k \mathfrak{D}_{K_1}(a_r^{d_r})$ ist bei $f(x)$ fixpunktfrei bis auf die Fixpunkte p und q ; der die Punkte $p, q_1, q_2, \dots, q_k, q$ verbindende Streckenzug \mathfrak{C} ist homotop zu seinem Bilde \mathfrak{C}' .

Beweis. Da p und q zu derselben Fixpunktklasse gehören, gibt es einen Weg \mathfrak{C}_0 von p nach q , der zu seinem Bilde \mathfrak{C}'_0 homotop ist. Diese Eigenschaft von \mathfrak{C}_0 bleibt erhalten bei Deformation von \mathfrak{C}_0 (unter Festhaltung der Endpunkte), da auch \mathfrak{C}'_0 dann nur eine Deformation erleidet. Also kann man \mathfrak{C}_0 ersetzen durch eine Kurve \mathfrak{C}_1 , die homotop zu \mathfrak{C}_0 ist und ganz in einem dreidimensional zusammenhängenden Gebiet \mathfrak{G}_0 verläuft. \mathfrak{C}_1 kann als Streckenzug gewählt werden. Durch eine kleine Verschiebung der Eckpunkte von \mathfrak{C}_1 kann man noch erreichen, daß \mathfrak{C}_1 doppelpunktfrei ist, die von p und q verschiedenen Fixpunkte von $f(x)$ und alle null- und eindimensionalen Zellen von K vermeidet und bis auf endlich viele Punkte im Innern von Grundsimplexen (von K) verläuft. Durch Unterteilung von K zu K_1 kann man außerdem erzielen, daß \mathfrak{C}_1 jedes Simplex von K_1 nur einmal (entweder in einem Punkt oder in einem offenen Streckenzug) trifft, daß die getroffenen Simplexe a samt ihrer Umgebung $\mathfrak{D}_{K_1}(a)$ frei von Fixpunkten außer p und q sind, und daß \mathfrak{C}_1 innerhalb von Grundsimplexen stets geradlinig verläuft. Die der Reihe nach getroffenen Grundsimplxe seien $a_0^{r_0}, a_1^{r_1}, \dots, a_k^{r_k}$. Wir erhalten \mathfrak{C} , indem wir jeden der (auf Nebensimplexen liegenden) Eckpunkte von \mathfrak{C}_1 neuerdings so verschieben, daß er im Innern der höchstdimensionalen gemeinsamen Randseite der beteiligten Grundsimplxe $a_{r-1}^{r_{r-1}}$ und $a_r^{r_r}$ liegt; damit haben wir die Punkte $q_r \in a_r^{d_r}$, und es ist $\overline{a_r^{d_r}} = \overline{a_{r-1}^{r_{r-1}}} \cap \overline{a_r^{r_r}}$ (vgl. Fig. 6). Damit ist alles bewiesen.

In Fig. 6 ist aus zeichentechnischen Gründen, im Widerspruch zu der Bedingung $d_r \geq 2$, $d_1 = d_2 = 1$, $r_0 = r_1 = 2$ gesetzt. — Im Falle (M) ist stets $r_r = d_r$, $d_r = d - 1$, wenn d die Dimension von \mathfrak{P} ist.

Das Gebiet $\mathfrak{G} = \sum \mathfrak{D}_{K_1}(a_r^{d_r})$ enthält auch die $\mathfrak{D}_{K_1}(a_r^{r_r})$. Nur in \mathfrak{G} werden wir (in Hilfssatz 14) die Funktion $f(x)$ abändern. — Auf \mathfrak{G} führen wir einen Parameter ξ ($0 \leq \xi \leq 1$) ein, der zur Bogenlänge proportional ist; es sei $c(\xi)$ die Parameterdarstellung von \mathfrak{G} , $c(0) = p$, $c(1) = q$.

Hilfssatz 12. Es gibt eine in $0 \leq \xi \leq 1$, $0 \leq \tau \leq 1$ stetige Funktion $h(\xi, \tau)$ mit Werten aus \mathfrak{P} und folgenden Eigenschaften:

$$h(\xi, 0) = c(\xi), \quad h(\xi, 1) = c'(\xi) \quad (= f(c(\xi))),$$

$$h(0, \tau) = p, \quad h(1, \tau) = q,$$

$$(63) \quad h(\xi, \tau) \neq c(\xi) \quad \text{für } \xi \neq 0, 1, \tau \neq 0.$$

Beweis. Alle Bedingungen bis auf (63) sind trivialerweise zu verwirklichen als Ausdruck der Tatsache, daß \mathfrak{G}' zu \mathfrak{G} homotop ist.

I. Es sei ε mit $0 < \varepsilon < \frac{1}{4}$ so gewählt, daß für $\xi \leq \varepsilon$ $c(\xi) \in a_0^{r_0}$, $c'(\xi) \in a_0^{r_0}$, und für $\xi \geq 1 - \varepsilon$ $c(\xi) \in a_k^{r_k}$, $c'(\xi) \in a_k^{r_k}$ ist (Fig. 6). Wir nehmen außerdem an, daß $c'(\varepsilon) \notin \mathfrak{G}$, $c'(1 - \varepsilon) \notin \mathfrak{G}$ ist; dies läßt sich nötigenfalls erzwingen durch eine beliebig kleine Deformation von $f(x)$ in der Umgebung von $c(\varepsilon)$ und $c(1 - \varepsilon)$, die unsere Voraussetzungen nicht stört. Dann setzen wir für $0 \leq \xi \leq \varepsilon$ und für $1 - \varepsilon \leq \xi \leq 1$

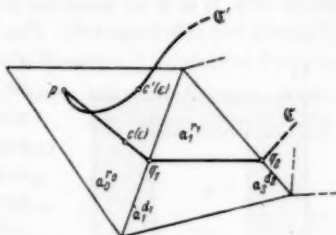


Fig. 6.

$$h(\xi, \tau) = (1 - \tau) c(\xi) + \tau c'(\xi) \quad (34).$$

Damit ist für $\xi \leq \varepsilon$ und $\xi \geq 1 - \varepsilon$ (sowie für $\tau = 0, 1$) $h(\xi, \tau)$ im Einklang mit (63) erklärt.

II. Wir verbinden die Punkte $h(\varepsilon, \varepsilon)$ und $h(1 - \varepsilon, \varepsilon)$ durch einen Streckenzug \mathfrak{G}_ε . Dieser soll dieselben Simplexe ($a_0^{r_0}$, $a_1^{r_1}$, ..., $a_k^{r_k}$) wie \mathfrak{G} durchlaufen und zwar die $a_r^{d_r}$ nur in je einem Punkte treffen, soll aber zu \mathfrak{G} punktfremd sein, was offenbar erreicht werden kann. Für \mathfrak{G}_ε wählen wir eine Parameterdarstellung $h(\xi, \varepsilon)$ ($\varepsilon \leq \xi \leq 1 - \varepsilon$), so daß stets $h(\xi, 0) (= c(\xi))$ und $h(\xi, \varepsilon)$ den gleichen Träger in K_1 haben. Dann sei für $\varepsilon \leq \xi \leq 1 - \varepsilon$, $0 \leq \tau \leq \varepsilon$

$$h(\xi, \tau) = \left(1 - \frac{\tau}{\varepsilon}\right) h(\xi, 0) + \frac{\tau}{\varepsilon} h(\xi, \varepsilon).$$

³⁴⁾ Deutsche Buchstaben bedeuten die entsprechenden Ortsvektoren, vgl. S. 549.

III. Für $\varepsilon \leq \xi \leq 1 - \varepsilon$ ist $h(\xi, 1) = c'(\xi) \neq c(\xi)$; es sei dort

$$(64) \quad \varrho(c(\xi), c'(\xi)) > \delta > 0.$$

Wir wenden Hilfssatz 1 (aus W.) an; $a(x, y, \tau)$ und $\alpha(\varepsilon)$ mögen die dortige Bedeutung haben. Wir wählen ε_1 mit $0 < \varepsilon_1 < \varepsilon$ so, daß

$$\varrho(h(\varepsilon, 1 - \varepsilon_1), c'(\varepsilon)) < \alpha(\delta), \quad \varrho(h(1 - \varepsilon, 1 - \varepsilon_1), c'(1 - \varepsilon)) < \alpha(\delta)$$

ist. Die Punkte $h(\varepsilon, 1 - \varepsilon_1)$ und $h(1 - \varepsilon, 1 - \varepsilon_1)$ verbinden wir durch einen Streckenzug $\mathfrak{C}^* = \{c^*(\xi) \mid \varepsilon \leq \xi \leq 1 - \varepsilon\}$, der zu \mathfrak{C} punktfremd ist, so daß für $\varepsilon \leq \xi \leq 1 - \varepsilon$

$$\varrho(c^*(\xi), c'(\xi)) < \alpha(\delta)$$

gilt³⁵⁾. Auf \mathfrak{C}^* soll ξ stückweise zur Bogenlänge proportional sein. Dann setzen wir

$$h(\xi, \tau) = a\left(c^*(\xi), c'(\xi), \frac{\tau - 1 + \varepsilon_1}{\varepsilon_1}\right) \quad \text{für } \varepsilon \leq \xi \leq 1 - \varepsilon, \quad 1 - \varepsilon_1 \leq \tau \leq 1.$$

Es ist dann nach W., Hilfssatz 1 $\varrho(h(\xi, \tau), c'(\xi)) < \delta$, daher nach (64) auch hier (63) gültig.

IV. Die gewünschte stetige Abbildung des Einheitsquadrates der (ξ, τ) -Ebene (Fig. 7) in \mathfrak{P} ist nunmehr für die in der Figur schraffierten Teile unter Wahrung von (63) hergestellt. Das Bild des Randes des durch $\varepsilon < \xi < 1 - \varepsilon$,

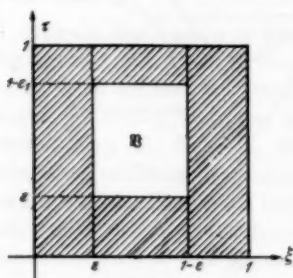


Fig. 7.

$\varepsilon < \tau < 1 - \varepsilon_1$ gegebenen Vierecks \mathfrak{B} ist nach Konstruktion ein zu \mathfrak{C} fremdes geschlossenes Polygon \mathfrak{B} . Dieses ist in \mathfrak{P} zusammenziehbar, weil der Rand des Quadrates in Fig. 7 durch $h(\xi, \tau)$ auf den nullhomotopen Weg $\mathfrak{C}\mathfrak{C}'^{-1}$ abgebildet wird. Wir wählen eine Unterteilung K_2 von K_1 , die eine Zerlegung von \mathfrak{B} als Teilkomplex enthält, jedoch so, daß \mathfrak{C} zu jeder null- oder eindimensionalen Zelle von K_2 fremd ist und mit jeder zweidimensionalen höchstens einen Punkt gemeinsam hat. Es ist dann

\mathfrak{B} zusammenziehbar in $|K_2^2|$, wo K_2^2 der maximale zweidimensionale Teilkomplex von K_2 ist; es gibt also eine Simplicialabbildung $h_1(\xi, \tau)$ von \mathfrak{B} in $|K_2^2|$, die sich stetig an die auf dem schraffierten Bereich erklärte Abbildung $h(\xi, \tau)$ anschließt. Die Bildmenge $h_1(\mathfrak{B})$ hat also mit \mathfrak{C} nur endlich viele Punkte gemeinsam³⁶⁾. Auch in \mathfrak{B} gibt es nur endlich viele Zahlenpaare (ξ, τ) mit $h(\xi, \tau) \in \mathfrak{C}$.

³⁵⁾ Man konstruiere den Weg mittels eines Sehnepolygons von \mathfrak{C}' , dessen Ecken man nachträglich, um \mathfrak{C} nicht zu treffen, etwas verschieben kann.

³⁶⁾ An dieser Stelle wird die Dimensionsvoraussetzung für $\mathfrak{P} = D$, S. 569 – ausgenutzt.

V. Ist $x_0 \in a^2 \in K_2$, $x_0 = c(\xi_0) = h_1(\xi_0, \tau_0)$, so sei v ein Vektor in Richtung einer x_0 enthaltenden oder von x_0 berandeten Strecke aus \mathbb{C} . Durch den Ansatz

$$h(\xi, \tau) = h_1(\xi, \tau) + \lambda(\xi, \tau) v,$$

wo λ nur in einer kleinen Umgebung von (ξ_0, τ_0) von Null verschieden sein soll, erhält man als Schnittpunkt von $h(\mathbb{B})$ mit \mathbb{C} an Stelle von x_0 einen anderen Punkt $h(\xi_0, \tau_0) \neq c(\xi_0)$. In dieser Weise kann man alle Punkte, an denen (63) verletzt ist, beseitigen. Damit ist Hilfssatz 12 bewiesen.

Hilfssatz 13. Es gibt eine für $x \in \mathbb{G}$ erklärte Abbildung \tilde{x} mit folgenden Eigenschaften:

$$\tilde{x} \in \mathbb{C}; \quad \tilde{x} = x \text{ für } x \in \mathbb{C}; \quad \tilde{x} \in Tr_{K_1}(x); \quad \tilde{y} = \tilde{x} \text{ für } y \in x \tilde{x}^{37}.$$

Beweis. Wir unterscheiden drei Fälle:

I. $x \in \mathfrak{D}_{K_1}(a_r^{d_i})$, $x \notin a_{r-1}^{r-1}$, $x \in a_r^{r*}$;

II. $x \in a_r^{r*}$, $r \neq 0, k$;

III. $x \in a_0^{r_0}$ oder $x \in a_k^{r_k}$.

Im Falle I sei $\tilde{x} = q_r$.

Im Falle II sei $x \notin \mathbb{C}$; \mathfrak{H}^2 sei die durch die Strecke q, q_{r+1} und den Punkt x bestimmte Halbebene. \mathfrak{H}^2 schneidet a_r^{r*} in einem konvexen Polygon, dessen Ecken der Reihe nach etwa $s, q_r, q_{r+1}, t, \dots$ sein mögen (Fig. 8). Wir erklären x durch die Festsetzung, daß die drei die Strecken sq_r , tq_{r+1} und $x\tilde{x}$ tragenden Geraden entweder durch einen Punkt z gehen oder parallel sein sollen.

Im Falle III, etwa für $x \in a_0^{r_0}$, verfahren wir analog wie in Fall II, indem wir q_0 als zweiten Schnittpunkt von $a_0^{r_0}$ mit der durch p und q_1 gehenden Geraden definieren. Ordnet aber das unter II angegebene Verfahren dem Punkt x einen nicht auf \mathbb{C} liegenden Punkt zu, so setzen wir statt dessen $\tilde{x} = p$. Für $x \in a_k^{r_k}$ verläuft die Konstruktion entsprechend.

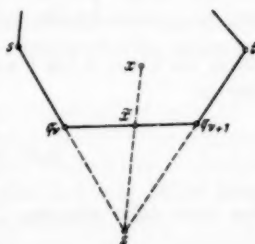


Fig. 8.

Daß die so erklärte Funktion x stetig ist und auch die weiteren gewünschten Eigenschaften hat, ist leicht zu erkennen. —

Es sei $\varkappa > 0$ so gewählt, daß der Abstand von $\mathbb{B} - \mathbb{G}$ und der Punktmenge $[\varrho(x, \tilde{x}) < \varkappa]$ ungleich Null ist. Mit \mathcal{U}_ε ($0 < \varepsilon < \varkappa$) bezeichnen wir die Menge der $x \in \mathbb{G}$ mit $\varrho(x, \tilde{x}) < \varepsilon$; \mathcal{U}_ε ist ein dreidimensional zusammen-

³⁷⁾ Die Abbildung \tilde{x} ist eine „Retraktion“; vgl. AH., S. 342.

hängendes Gebiet. $\xi(x)$ sei für $x \in \mathfrak{G}$ die durch $\tilde{x} = c(\xi(x))$ bestimmte Funktion; $\xi(x)$ ist stetig.

Hilfssatz 14. *Es gibt eine zu $f(x)$ homotope Selbstabbildung $g_1(x)$ und ein Gebiet $\mathfrak{U} \subset \mathfrak{G}$ mit folgenden Eigenschaften: für $x \in \mathfrak{P} - \mathfrak{G}$ ist $g_1(x) = f(x)$; für $x \in \mathfrak{G} - \mathfrak{U}$ ist $g_1(x) \neq x$; für $x \in \mathfrak{U}$ ist $g_1(x) \in \mathfrak{P}_K(x)$.*

Beweis. Es sei ε eine Zahl mit $0 < \varepsilon < \kappa$, über die noch verfügt wird. Wir definieren

$$(v5) \quad g_1(x) = \begin{cases} h\left(\xi(x), \frac{2}{\varepsilon} \varrho(x, \tilde{x})\right) & \text{für } x \in \mathfrak{U}_{\frac{\varepsilon}{2}}, \\ f\left(\left(2 - \frac{2}{\varepsilon} \varrho(x, \tilde{x})\right) \tilde{x} + \left(\frac{2}{\varepsilon} \varrho(x, \tilde{x}) - 1\right) x\right) & \text{für } x \in \mathfrak{U}_{\frac{\varepsilon}{2}} - \mathfrak{U}_{\frac{\varepsilon}{2}}, \\ f(x) & \text{für } x \in \mathfrak{P} - \mathfrak{U}_{\frac{\varepsilon}{2}}. \end{cases}$$

Man überzeugt sich zunächst: $g_1(x)$ ist für $x \in \mathfrak{P}$ definiert. $g_1(x)$ ist auf den Mengen $\mathfrak{U}_{\frac{\varepsilon}{2}}$, $\mathfrak{U}_{\frac{\varepsilon}{2}} - \mathfrak{U}_{\frac{\varepsilon}{2}}$ und $\mathfrak{P} - \mathfrak{U}_{\frac{\varepsilon}{2}}$ stetig. Auf den Begrenzungen dieser

Mengen, d. h. für $\varrho(x, \tilde{x}) = \frac{\varepsilon}{2}$ bzw. $\varrho(x, \tilde{x}) = \varepsilon$, stimmen der erste und zweite bzw. der zweite und dritte Ausdruck für $g_1(x)$ überein; $g_1(x)$ ist damit auf \mathfrak{P} stetig.

Um die Homotopie von g_1 mit f nachzuweisen, führen wir die Deformation von g_1 in f explizit aus: es sei für $0 \leq \sigma \leq 1$

$$g_1(\sigma; x) = \begin{cases} h\left(\xi(x), \sigma + (1 - \sigma) \frac{2}{\varepsilon} \varrho(x, \tilde{x})\right) & \text{für } x \in \mathfrak{U}_{\frac{\varepsilon}{2}}, \\ g_1(x) & \text{für } x \in \mathfrak{P} - \mathfrak{U}_{\frac{\varepsilon}{2}}. \end{cases}$$

Indem wir σ von 0 bis 1 wachsen lassen, geht dann $g_1(0; x) = g_1(x)$ stetig über in

$$g_1(1; x) = \begin{cases} f(\tilde{x}) & \text{auf } \mathfrak{U}_{\frac{\varepsilon}{2}}, \\ g_1(x) & \text{auf } \mathfrak{P} - \mathfrak{U}_{\frac{\varepsilon}{2}}. \end{cases}$$

Man kann dafür schreiben

$$g_1(1; x) = \begin{cases} f(\lambda \tilde{x} + (1 - \lambda) x) & \text{auf } \mathfrak{U}_{\frac{\varepsilon}{2}}, \\ f(x) & \text{auf } \mathfrak{P} - \mathfrak{U}_{\frac{\varepsilon}{2}}; \end{cases}$$

hierbei ist $\lambda = \lambda(x)$ eine wohlbestimmte auf $\mathfrak{U}_{\frac{\varepsilon}{2}}$ stetige Ortsfunktion mit $0 \leq \lambda \leq 1$. Setzen wir nun für $1 \leq \sigma \leq 2$

$$g_1(\sigma; x) = \begin{cases} f(\lambda(2 - \sigma) \tilde{x} + (1 - 2\lambda + \lambda\sigma) x) & \text{auf } \mathfrak{U}_{\frac{\varepsilon}{2}}, \\ f(x) & \text{auf } \mathfrak{P} - \mathfrak{U}_{\frac{\varepsilon}{2}}, \end{cases}$$

so haben wir den Übergang von $g_1(1; x)$ nach $g_1(2; x) = f(x)$. g_1 ist also homotop zu f .

Es sei nun η die nach AH. VIII, § 2, 1 existierende (dort mit τ bezeichnete) Zahl, so daß jede Menge aus \mathfrak{B} mit einem Durchmesser kleiner als η ganz in einem Stern $\mathfrak{D}_K(u_i)$ (u_i eine Ecke von K) enthalten ist. Es gilt dann auch:

$$(66) \quad \text{aus } \varrho(x, y) < \eta \text{ folgt } y \in \mathfrak{B}_K(x).$$

Da für $\xi = 0$, $\xi = 1$ und $\tau = 0$ $\varrho(h(\xi, \tau), c(\xi)) = 0$ ist, läßt sich ferner ein δ mit $0 < \delta < \frac{1}{2}$ bestimmen, so daß

$$(67) \quad \varrho(h(\xi, \tau), c(\xi)) < \frac{\eta}{3} \\ \text{für } \xi < \delta, \text{ für } \xi > 1 - \delta \text{ und für } \tau < \delta$$

ist. Weiter sei $\zeta > 0$ so gewählt, daß

$$(68) \quad \varrho(h(\xi, \tau), c(\xi)) > \zeta \quad \text{für } \delta \leq \xi \leq 1 - \delta, \quad \tau \geq \delta$$

ist; für die abgeschlossene Menge $[\delta \leq \xi \leq 1 - \delta, \tau \geq \delta]$ ist ja $h(\xi, \tau) \neq c(\xi)$ (Hilfssatz 12). Schließlich bestimmen wir ε mit

$$(69) \quad 0 < \varepsilon < \min\left(\frac{\eta}{3}, \frac{\zeta}{2}\right)$$

so, daß

$$(70) \quad \varrho(x', y') < \min\left(\frac{\eta}{3}, \frac{\zeta}{2}\right) \quad \text{für } \varrho(x, y) < \varepsilon$$

ist. Damit ist g_1 dann durch (65) festgelegt. \mathfrak{U} bestehe aus den Punkten $x \in \mathfrak{U}_i$ mit $\xi(x) < \delta$ oder $\xi(x) > 1 - \delta$ und den Punkten $x \in \mathfrak{U}_{i\delta}$; \mathfrak{U} ist offen und, da es \mathfrak{C} und mit jedem x die Strecke xx enthält, zusammenhängend, also ein Gebiet. Es ist $\mathfrak{U} \subset \mathfrak{U}_i$. Da g_1 nur auf \mathfrak{U}_i von f verschieden ist, haben wir nur noch zu zeigen:

a) auf $\mathfrak{U}_i - \mathfrak{U}$ ist $g_1(x) \neq x$;

b) auf \mathfrak{U} ist $g_1(x) \in \mathfrak{B}_K(x)$.

ad a). Es sei $\frac{\varepsilon\delta}{2} \leq \varrho(x, \tilde{x}) < \varepsilon$ und $\delta \leq \xi(x) \leq 1 - \delta$. Es ist dann nach (65) entweder $\varrho(x, \tilde{x}) < \frac{\varepsilon}{2}$ und

$$\varrho(g_1(x), x) = \varrho(h(\xi(x), \tau), x) \geq \varrho(h(\xi(x), \tau), c(\xi(x))) - \varrho(x, \tilde{x})$$

mit $\tau = \frac{2}{\varepsilon} \varrho(x, \tilde{x}) \geq \delta$, also wegen (68) und (69)

$$\varrho(g_1(x), x) > \zeta - \varepsilon > \frac{\zeta}{2} > 0,$$

oder $\varrho(x, \tilde{x}) \geq \frac{\varepsilon}{2}$ und

$$\begin{aligned} \varrho(g_1(x), x) &= \varrho((\lambda \tilde{x} + (1 - \lambda)x)', x) \\ &\geq -\varrho((\lambda \tilde{x} + (1 - \lambda)x)', \tilde{x}) + \varrho(h(\xi(x), 1), c(\xi)) - \varrho(x, \tilde{x}) \end{aligned}$$

mit $0 \leq \lambda \leq 1$; also, da $\varrho(\lambda \tilde{x} + (1 - \lambda)x, \tilde{x}) \leq \varrho(x, \tilde{x}) < \varepsilon$ ist, nach (68), (69), (70)

$$\varrho(g_1(x), x) > -\frac{\varepsilon}{2} + \zeta - \varepsilon > \frac{\varepsilon}{2} - \frac{\varepsilon}{2} = 0,$$

also in jedem Falle $g_1(x) \neq x$.

ad b). Es genügt zu zeigen: $\varrho(g_1(x), x) < \eta$ für $x \in \mathcal{U}$. Ist $x \in \mathcal{U}$, und $\xi(x) < \delta$ oder $\xi(x) > 1 - \delta$, so ist entweder nach (67) und (69)

$$\begin{aligned}\varrho(g_1(x), x) &= \varrho(h(\xi(x), \tau), x) \\ &\leq \varrho(h(\xi(x), \tau), c(\xi(x))) + \varrho(x, \tilde{x}) < \frac{\eta}{3} + \varepsilon < \eta,\end{aligned}$$

oder nach (67) und (70)

$$\begin{aligned}\varrho(g_1(x), x) &= \varrho((\lambda \tilde{x} + (1 - \lambda)x)', x) \\ &\leq \varrho((\lambda \tilde{x} + (1 - \lambda)x)', \tilde{x}') + \varrho(h(\xi(x), 1), c(\xi(x))) + \varrho(x, \tilde{x}) \\ &\leq \frac{\eta}{3} + \frac{\eta}{3} + \varepsilon < \eta.\end{aligned}$$

Ist ferner $x \in \mathcal{U}_{\frac{\delta}{2}}$, also auch $x \in \mathcal{U}_{\frac{\varepsilon}{2}}$, so ist $\varrho(g_1(x), x) = \varrho(h(\xi(x), \tau), x)$ mit

$$\tau = \frac{2}{\varepsilon} \varrho(x, \tilde{x}) < \delta, \text{ also nach (65), (67), (70)}$$

$$\varrho(g_1(x), x) \leq \varrho(h(\xi(x), \tau), c(\xi(x))) + \varrho(x, \tilde{x}) < \frac{\eta}{3} + \varepsilon < \eta.$$

Damit ist Hilfssatz 14 bewiesen.

Hilfssatz 15. Die auf S. 570 bezeichnete Abbildung f läßt sich stetig in eine solche Abbildung g deformieren, die in einer Umgebung der von p und q verschiedenen Fixpunkte von $f(x)$ mit f übereinstimmt und außer ihnen nur noch einen und zwar einen regulären Fixpunkt hat, der wieder in einem Grundsimpler von K liegt.

Beweis. Wir gehen von der in Hilfssatz 14 konstruierten Funktion $g_1(x)$ aus, die wir nur noch innerhalb \mathcal{U} abzuändern haben, so daß dort nur ein Fixpunkt bleibt. \mathcal{M} sei eine abgeschlossene, zusammenhängende Teilmenge von \mathcal{U} , die alle Fixpunkte aus $\mathcal{U}^{(20)}$ enthält; \mathcal{M} habe von $\mathcal{P} - \mathcal{U}$ den Abstand $\varepsilon > 0$. K_2 sei eine Unterteilung von K_1 , bei der alle Simplexdurchmesser kleiner als ε sind (eine ε -Zerlegung, vgl. W., S. 663). Ω sei die Vereinigung aller abgeschlossenen Grundsimpler von K_2 , die von \mathcal{M} den Abstand Null haben; es ist $\mathcal{M} \subset \Omega \subset \mathcal{U}$. \mathcal{N} sei das Innere, $\bar{\Omega}$ die Begrenzung von Ω , es sei $\mathcal{R} = \mathcal{P} - \mathcal{N}$. Jedes Simplex von \mathcal{R} hat von \mathcal{M} einen von Null verschiedenen Abstand; die Punkte von \mathcal{M} und insbesondere die Fixpunkte aus \mathcal{U} liegen daher in \mathcal{N} . \mathcal{N} ist nach Konstruktion zusammenhängend, und zwar als Teilgebiet von \mathcal{U} dreidimensional, also erst recht zweidimensional zusammenhängend. Die Abbildung $g_1(x)$ ist auf $\bar{\Omega}$ fixpunktfrei, und es ist dort $g_1(x)$

²⁰⁾ Also insbesondere alle Punkte der Kurve \mathcal{C} .

$\in \mathfrak{B}_K(x)$. Es sei nun $g(x) = g_1(x)$ für $x \in \mathfrak{U}$; nach Satz 1 setzen wir $g(x)$ über \mathfrak{U} fort, so daß (3) gilt und nur ein Fixpunkt in einem Grundsimplex von K_2 entsteht. Wir sind fertig, sobald noch gezeigt ist, daß g zu g_1 homotop ist.

$\delta > 0$ sei der Abstand von Ω und $\mathfrak{P} - \mathfrak{U}$. Aus (3) folgt nach S. 552f. leicht die Existenz eines Weges von x über $h(x)$ nach $g(x)$:

$$a(x, g(x), \tau) \text{ stetig in } x \text{ und } \tau \quad (0 \leq \tau \leq 1),$$

$$a(x, g(x), 0) = x, a(x, g(x), 1) = g(x)$$

(vgl. auch W., Hilfssatz 1 und AH., S. 343). Ist $\zeta(x)$ der Abstand von Ω und x , so ist für $\zeta(x) < \delta$ $g_1(x) \in \mathfrak{B}_K(x)$, $g(x) \in \mathfrak{B}_K(x)$ und für $\zeta(x) > 0$ $g_1(x) = g(x)$. Man kommt nun von $g(x)$ über

$$g_\sigma(x) = a\left(x, g(x), 1 - \sigma \max\left(1 - \frac{\zeta(x)}{\delta}, 0\right)\right) \quad (0 \leq \sigma \leq 1)$$

stetig zu einer Funktion

$$g^*(x) = \begin{cases} a\left(x, g(x), \frac{\zeta(x)}{\delta}\right) & \text{wenn } \zeta(x) \leq \delta, \\ g(x) & \text{wenn } \zeta(x) \geq \delta \end{cases}$$

und analog von $g_1(x)$ aus zu der damit identischen Funktion

$$g_1^*(x) = \begin{cases} a\left(x, g_1(x), \frac{\zeta(x)}{\delta}\right) & \text{wenn } \zeta(x) \leq \delta, \\ g_1(x) & \text{wenn } \zeta(x) \geq \delta. \end{cases}$$

Nach diesen Vorbereitungen können wir den Hauptsatz dieses Paragraphen fast ohne Beweis aussprechen.

Satz 3. Das Polyeder \mathfrak{P} habe die auf S. 569 erklärte Eigenschaft D. Dann enthält jede Abbildungsklasse \mathfrak{f} von \mathfrak{P} in sich Abbildungen, bei denen jede wesentliche Fixpunktklasse durch einen einzigen Fixpunkt vertreten ist und unwesentliche Fixpunktklassen nicht auftreten.

Beweis. $f(x)$ sei eine Abbildung aus \mathfrak{f} mit nur regulären Fixpunkten und übrigens mit möglichst kleiner geometrischer Fixpunktzahl. Es kann bei $f(x)$ insbesondere kein Fixpunkt der Vielfachheit Null auftreten. Gäbe es mehr Fixpunkte bei f , als die Zahl μ der wesentlichen Fixpunktklassen angibt, so müßte eine Fixpunktklasse zwei reguläre Fixpunkte enthalten, und die geometrische Fixpunktzahl ließe sich — gegen die Voraussetzung — nach Hilfssatz 15 noch erniedrigen unter Aufrechterhaltung der Regularitätsbedingung. Nach der schon in W. bewiesenen Ungleichung $m \geq \mu$ kann ferner auch eine Abbildung mit nichtregulären Fixpunkten deren nicht weniger als μ haben.

Der wesentliche Inhalt von Satz 3 kann so ausgedrückt werden (vgl. die Bemerkung nach Satz 2, S. 569): Für jede Abbildungsklasse bei den der Bedingung D genügenden Polyedern gilt $m = \mu$.

Über eine Eigenschaft von Systemen linearer wohlgeordneter Mengen.

Von

Geörg Kurepa in Agram (Kroatien).

Am anderen Ort habe ich folgendes Problem gestellt¹⁾: Sei F ein beliebiges abzählbares System wohlgeordneter Untermengen²⁾ der geordneten Menge der rationalen Zahlen; enthält dann F ein *gleichmächtiges* Untersystem, das so beschaffen ist, daß von je zwei seiner Elemente keines ein Anfangsstück vom anderen ist?

Auf die vorstehende Frage werden wir *bejahend* antworten und beweisen, daß ein analoger Satz sogar für Systeme wohlgeordneter *linearer* Mengen zutrifft, ja sogar für Systeme wohlgeordneter Untermengen einer beliebigen geordneten *separablen* Menge³⁾. Ob dies noch wahr ist für Systeme wohlgeordneter Mengen einer beliebigen geordneten mit *Suslinscher Eigenschaft* behafteten Menge⁴⁾, bleibt noch offen (und ist gleichbedeutend mit dem Suslinschen Problem: ob jede stetige, geordnete mit Suslinscher Eigenschaft behaftete Menge zu einer *linearen* Menge ähnlich ist).

Die Begriffe über die sogenannte *verzweigte Tafel*⁵⁾ (s. u.) passen sich gut der Sache an, und der Inhalt des Aufsatzes gipfelt in der Tatsache, daß *jede verzweigte Tafel, in der eine reelle wachsende Funktion existiert, normal ist* (s. u.).

1. *Verzweigte Tafel*. Eine teilweise geordnete Menge T (bezüglich einer ordnenden Relation $<$) ist eine *verzweigte Tafel*⁵⁾, im folgenden v. T., oder T (bezüglich $<$), wenn für jedes $a \in T$ die Menge

$$(1) \quad (., a)_T$$

aller $x \in T$ mit $x < a$ *wohlgeordnet* ist (bezüglich $<$).

¹⁾ Georges Kurepa, Ensembles ordonnés et ramifiés, Thèse, Paris 1935; auch in Publ. Math. Univ. Belgrade IV (1935), S. 1–135, 3, 134, Fußnote.

²⁾ Bezüglich der Terminologie über die Mengenlehre s. Felix Hausdorff, Mengenlehre, 2. Auflage. Berlin-Leipzig 1927.

³⁾ Nach Fréchet heißt eine geordnete Menge *separabel*, wenn sie eine höchstens abzählbare überall dichte Teilmenge enthält.

⁴⁾ Eine geordnete Menge besitzt die *Suslinsche Eigenschaft*, wenn jedes System disjunkter nichtleerer offener Intervalle der Menge höchstens abzählbar ist.

⁵⁾ Französisch: tableaux ramifiés [l. c.¹⁾, S. 73].

Für jede Ordnungszahl α werden wir durch

$$(2) \quad R_\alpha T$$

die Menge aller $x \in T$ bezeichnen, für welche die wohlgeordnete Menge $(., x)_T$ vom Typus α ist; die Menge (2) heiße α -Klasse von T (bezüglich $<$). Die erste Ordnungszahl α , für welche $R_\alpha T$ leer ist, heiße *Rang* von T und werde durch

$$(3) \quad \gamma T$$

bezeichnet. So haben wir also die wohldefinierte Folge der Klassen von T

$$(4) \quad R_0 T, R_1 T, \dots, R_\alpha T, \dots \quad (\alpha < \gamma T),$$

und eine Klasse $R_\alpha T$ ist von 1. oder 2. Art, je nachdem die Zahl α von 1. oder 2. Art ist.

Folgende Tatsachen sind fast selbstverständlich:

a) Jeder Punkt von T gehört einer und nur einer Klasse an; daher die Redeweise: ein Punkt von T ist von 1. oder 2. Art, je nachdem ob die ihn enthaltende Klasse von 1. oder 2. Art ist.

b) $T = \sum_a R_a T$, ($\alpha < \gamma T$); $T \supset \sum_a R_a T^a$, ($\alpha < \beta$), für jedes $\beta < \gamma T$.

c) Die Elemente jeder Klasse sind unvergleichbar, d. h. wenn a, b zwei verschiedene, zu derselben Klasse gehörige Punkte sind, dann ist weder $a < b$ noch $b < a$.

d) Wenn $\alpha < \beta < \gamma T$ und $b \in R_\beta T$, dann enthält die Klasse $R_\alpha T$ ein einziges Element a mit $a < b$.

2. *Normale Tafel.* Ein T ist (bezüglich $<$) *degeneriert*, wenn für jedes $a \in T$ die Menge aller mit a vergleichbaren Punkte von T *geordnet* (also auch *wohlgeordnet*) ist. Z. B. jede Klasse von T ist degeneriert. Eine v. T. ist *normal*, wenn sie endlich ist oder eine *gleichmächtige degenerierte* Untermenge enthält.

Die Frage, ob jede v. T. normal ist, bleibt offen; jedenfalls beweist man folgende Hilfssätze:

Hilfssatz 1. Jede höchstens abzählbare v. T. ist normal⁷⁾.

Hilfssatz 2. Wenn die Potenz von T größer ist als die von γT , dann ist T normal⁸⁾.

3. *Die Tafel E_w* (hierbei bedeute E eine beliebige geordnete Menge). Wenn E eine geordnete Menge ist, so werden wir wie Hartogs⁹⁾ durch

$$(5) \quad E_w$$

⁶⁾ Das Inklusionszeichen \supset schließt das Gleichheitszeichen $=$ aus.

⁷⁾ l. c.¹⁾, § 11, Nr. 2a.

⁸⁾ l. c.¹⁾, § 11, Nr. 2 u. 3.

⁹⁾ F. Hartogs, Über das Problem der Wohlordnung. Math. Annalen 76 (1915), S. 438–443.

das durch die Relation $<$ teilweise geordnete System aller wohlgeordneten Teilmengen von E bezeichnen. Dabei bedeute $A < B$, daß die wohlgeordnete Menge A ein Anfangsstück von der wohlgeordneten Menge B ist, ohne daß B ein Anfangsstück von A sei.

E_w ist eine bezüglich $<$ verzweigte Tafel, und für jedes α besteht $R_\alpha E_w$ aus allen wohlgeordneten Teilmengen von T , deren Ordnungstypus α ist; insbesondere besteht $R_0 E_w$ aus der Nullmenge.

Man beweist leicht, daß E_w normal ist; die schwere Frage aber, ob für jedes E auch jedes Teilsystem von E_w normal ist, bleibt unentschieden.

Im folgenden werden wir die folgenden vier Fälle behandeln:

Erster Fall: E ist die geordnete Menge der rationalen Zahlen.

Zweiter Fall: E ist die geordnete Menge der reellen Zahlen.

Dritter Fall: E ist eine beliebige separable geordnete Menge³⁾.

Vierter Fall: E ist eine beliebige geordnete, mit Suslinscher Eigenschaft behaftete Menge⁴⁾.

4. Erster Fall. Die Tafel E_w (E = geordnete Menge aller rationalen Zahlen) [vgl. (5)].

Satz 1. Jede Tafel $T \subseteq E_w$ ist normal; insbesondere enthält jedes abzählbare unendliche System T wohlgeordneter Teilmengen der Menge der rationalen Zahlen ein gleichmächtiges Teilsystem, in dem kein Element ein Anfangsstück von einem anderen ist.

Beweis. Nun sei $T \subseteq E_w$; nach Hilfssatz 1 kann man annehmen $pT \geq p\Omega^{10}$, nach dem Hilfssatz 2 können wir, wegen $\gamma E_w \leq \Omega$, annehmen $pT = p\Omega$. Weiter kann man supponieren, daß jede Klasse $R_\alpha T$ höchstens abzählbar ist (sonst hätte T dieselbe Potenz wie eine seiner Klassen, womit alles bewiesen wäre).

Wegen $\gamma T \leq \Omega$, $pR_\alpha T \leq p\epsilon_\alpha^{10}$ ($\alpha < \Omega$) haben wir also

$$\gamma T = \Omega, \quad pR_\alpha T \leq p\epsilon_\alpha, \quad pT = p\Omega^{10}.$$

Es seien nun:

$$(6) \quad A_0, A_1, \dots, A_\xi, \dots, \quad (\xi < \Omega)$$

die transfinite Folge aller Elemente von T , und

$$(7) \quad A = \{a_0, a_1, \dots, a_\xi, \dots\}_{\xi < \omega}$$

irgend ein Element von T (also auch von E_w); A ist also eine wohlgeordnete Menge von rationalen Zahlen vom Typus α . Für jede rationale Zahl r werde die Zahl

$$(8) \quad \varphi(r, A)$$

¹⁰⁾ pX bedeutet die Potenz von X ; insbesondere also $p\Omega = \aleph_1$, $p\omega = \aleph_0$.

folgendermaßen definiert:

$\varphi(r, A) = -1$, wenn r nicht in A enthalten ist;

$\varphi(r, A) = \beta$, wenn in der Darstellung (7) gerade $a_\beta = r$. Weiter setzen wir

$$(9) \quad \varphi(r, T) = \sup_A \varphi(r, A), \quad (A \varepsilon T).$$

Man hat $-1 \leq \varphi(r, T) \leq \Omega$.

a) Es gibt mindestens eine rationale Zahl r_0 mit $\varphi(r_0, T) = \Omega$. Wäre nämlich $\varphi(r, T) < \Omega$ für jedes r , so wäre auch $\delta < \Omega$, wo $\delta = \sup_r \varphi(r, T)$ und r die abzählbare Menge aller rationalen Zahlen durchläuft. Nun sei A ein Element von $R_{\delta+3}T$. Wegen $\gamma T = \Omega$ existiert A . Weil der Ordnungstypus von A mindestens $\delta + 3$ ist, so hätte man für die $(\delta + 2)$ -te Zahl r in der Menge A $\varphi(r, A) = \delta + 2$, also $\varphi(r, A) > \delta = \varphi(r, T)$ und damit einen Widerspruch.

Wenn so die Existenz der Zahl r_0 bewiesen ist, konstruieren wir die transfinite Folge

$$(10) \quad A^0, A^1, \dots, A^\xi, \dots, \quad (\xi < \Omega)$$

folgendermaßen:

A^0 sei das erste Glied in (6) mit $\varphi(r_0, A^0) \geq 0$; allgemein sei für jedes $0 < \xi < \Omega$ nun A^ξ das erste Glied in (6) mit $\varphi(r_0, A) > \sup_{\alpha < \xi} \varphi(r_0, A^\alpha)$; wegen

$\varphi(r_0, T) = \Omega$ ist für jedes $\xi < \Omega$ die Existenz von A^ξ gesichert, und es ist

$$(11) \quad \varphi(r_0, A^0) < \varphi(r_0, A^1) < \dots < \varphi(r_0, A^\xi) < \dots, \quad (\xi < \Omega).$$

b) Die Elemente der Folge (10) sind miteinander unvergleichbar, d. h. wenn $\alpha < \beta < \Omega$, so ist weder $A^\alpha < A^\beta$ noch $A^\beta < A^\alpha$.

Seien nämlich

$$A^\alpha = (a_0^\alpha, a_1^\alpha, \dots, a_\xi^\alpha, \dots)_{\xi < \gamma},$$

$$A^\beta = (a_0^\beta, a_1^\beta, \dots, a_\xi^\beta, \dots)_{\xi < \gamma},$$

die Darstellungen der Form (7) von A^α bzw. A^β . Definitionsgemäß ist

$$a_{\varphi(r_0, A^\alpha)}^\alpha = r_0 = a_{\varphi(r_0, A^\beta)}^\beta.$$

Wäre nun zum Beispiel $A^\alpha < A^\beta$, so wäre

$$a_{\varphi(r_0, A^\alpha)}^\beta = r_0, \quad \text{also} \quad a_{\varphi(r_0, A^\alpha)}^\beta = a_{\varphi(r_0, A^\beta)}^\beta,$$

was unmöglich ist, weil einerseits nach (11) $\varphi(r_0, A^\alpha) < \varphi(r_0, A^\beta)$ und andererseits wegen der Wohlordnung von A^β

$$a_{\varphi(r_0, A^\alpha)}^\beta < a_{\varphi(r_0, A^\beta)}^\beta.$$

Ebensowenig ist $A^\beta < A^\alpha$.

Zusammengefaßt können wir sagen: Die Menge der A^ξ ($\xi < \Omega$) in (10) ist eine unabzählbare normale Teiltabelle von $T \subseteq R_w$. Damit ist Satz 1 bewiesen.

5. Zweiter Fall. Tafel E_ω (E = Menge aller reellen Zahlen).

Satz 2. Jede Tafel $T \subseteq E_\omega$ ist normal. Jede unabzählbare Tafel $T \subseteq E_\omega$ enthält eine gleichmächtige Teiltafel von unvergleichbaren Elementen.

Setzen wir für jedes $A \in E_\omega$, welches von der Nullmenge verschieden ist,

$$f(A) = \sup_s x, \quad (x \in A),$$

so hat man $-\infty \leq f(A) \leq \infty$, und für jedes Tripel A, B, C von Elementen aus E_ω mit $A < B < C$ hat man $f(A) \leq f(B) \leq f(C)$ mit höchstens einem Gleichheitszeichen.

Wie im vorigen Falle können wir ohne Einschränkung voraussetzen, daß

$$(12) \quad pT = p\Omega, \quad \gamma T = \Omega, \quad pR_\alpha T \leq p\omega, \quad (\alpha < \Omega)^{(10)};$$

jede der beiden Mengen

$$T_1 = \sum_s R_{\alpha+1} T, \quad (\alpha < \Omega), \quad T_2 = T - T_1$$

genügt denselben Bedingungen, wie auch der folgenden:

Wenn A, B zwei Elemente von T_1 (bzw. T_2) sind, so daß $A < B$, dann ist $-\infty \leq f(A) < f(B) \leq \infty$. In bekannter Weise sagen wir dafür, daß $f(x)$ in T_1 (bzw. T_2) eine reelle wachsende Funktion ist.

Infolgedessen wird der Satz 2 durch folgenden schönen Satz erweitert:

Satz 3. Jede verzweigte Tafel T , in der eine reelle wachsende Funktion existiert, ist normal; falls T unabzählbar ist, hat T dieselbe Potenz wie eine seiner Untermengen, die kein Paar vergleichbarer Elemente enthält.

Beweis des Satzes 3. Sei nun

$$(13) \quad f(x)$$

eine reelle wachsende Funktion in T ; es ist $\gamma T \leq \Omega$ (gäbe es nämlich $a \in R_\alpha T$, so wäre die Menge aller Zahlen $f(x)$, ($x \in T$ mit $x < a$), eine unabzählbare wohlgeordnete lineare Menge, was nach einem Cantorsche Satz unmöglich ist). Wie oben können wir auch hier voraussetzen, daß (12) erfüllt ist.

6. Knoten von T . Für jedes $a \in T$ wird die Menge aller $x \in T$ mit $(., x)_T = (., a)_T$, (v. (1)) der Knoten von T , insbesondere der a -Knoten von T genannt. Z. B. ist die Klasse $R_0 T$ ein Knoten von T .

Jeder Knoten ist in einer bestimmten Klasse enthalten; dementsprechend wird er von 1. oder 2. Art benannt, je nachdem die zugehörige Klasse von 1. oder 2. Art ist. (Die 0-Klasse ist von 1. Art.)

7. Nun werden wir den Beweis des Satzes 3 auf den Fall zurückführen, wo jeder Knoten 2. Art einpunktig und wo die wachsende Funktion $f(x)$ in jedem Knoten 1. Art eineindeutig und rationalwertig ist.

Sei nämlich

$$(14) \quad T_0$$

das System aller Knoten 2. Art von T und aller Mengen, von denen jede aus einem einzigen Punkte 1. Art von T besteht; dabei wird T_0 durch die folgende Relation ϱ teilweise geordnet:

Wenn $A, B \in T_0$, so wird

$$(15) \quad A \varrho B$$

bedeuten, daß es ein $a \in A$ und ein $b \in B$ gibt mit $a < b^{11}$.

a) T_0 ist eine bezüglich ϱ verzweigte Tafel so, daß

$$\gamma T = \gamma T_0, \quad p R_a T_0 \leq p \omega, \quad (\alpha < \Omega), \quad p T = p T_0 = p \Omega.$$

b) Jeder Knoten 2. Art von T_0 ist einpunktig.

c) Sei X eine Teilmenge von T_0 mit lauter unvergleichbaren Elementen; wenn dann \bar{X} eine Menge bezeichnet, die gerade einen Punkt mit jedem Element von X gemeinsam hat — die Elemente von X sind doch gewisse Teilmengen von T — und keine weiteren Punkte hat, so ist \bar{X} eine bestimmte, mit X gleichmächtige Teilmenge von T selbst und enthält kein Paar [bezüglich $<^{11}$] vergleichbarer Punkte. Wenn daher T_0 normal ist, so ist es auch die Tafel T selbst.

Definieren wir noch eine wachsende Funktion $\varphi(x)$ in T_0 . Zuerst sei $F(a) = \sup_x f(x)$, ($x \in (\cdot, a)_{T_0}$), [vgl. (1)], wenn a ein Knoten 2. Art von T ist und $F(a) = f(a)$, wenn a aus einem Punkte 1. Art von T besteht. Dann ist $F(a)$ eine in T_0 wachsende reelle Funktion.

In jedem Knoten 1. Art K von T_0 definieren wir folgendermaßen eine *eindeutige rationalwertige Funktion* $\varphi_K(x)$, ($x \in K$): 1. $\varphi_K(x) < F(x)$, ($x \in K$), wenn $K = R_0 T_0$. 2. $F(k) < \varphi_K(x) < F(x)$, ($x \in K$), wenn $K \neq R_0 T_0$, worin k das wohlbestimmte Element von T_0 bezeichnet, welches dem Knoten K unmittelbar vorangeht.

Schließlich setzen wir für jedes $x \in T_0$

$$\varphi(x) = F(x) \quad \text{oder} \quad \varphi_{\xi}(x),$$

je nachdem ob x ein Element 2. oder 1. Art ist; im letzten Falle bezeichnet ξ den das Element x enthaltenden Knoten von T_0 .

$\varphi(x)$ ist eine in T_0 wachsende reelle in jedem Knoten von T_0 eindeutige Funktion; sie ist rationalwertig in jedem Knoten 1. Art von T_0 .

¹¹⁾ a, b sind also zwei Elemente von T ; $<$ ist die Relation, bezüglich derer T eine verzweigte Tafel ist.

8. T_0 ist normal. Für $a \in T_0$ sei $E(a)$ die Menge der Zahlen $\varphi(x)$, wo x die Menge aller Elemente 1. Art von T_0 durchläuft mit $x \in a$ [vgl. (15)]; außerdem solle $E(a)$ auch die Zahl $\varphi(a)$ enthalten, wenn a von 1. Art ist.

a) Für jedes $a \in T_0$ ist $E(a)$ eine wohlbestimmte wohlgeordnete nichtleere Menge rationaler Zahlen.

b) Wenn $a, b \in T_0$, dann ist $E(a) = E(b)$, $E(a) < E(b)$, $E(b) < E(a)$, je nachdem ob $a = b$, $a \in b$, $b \in a$.

c) Wenn a, b zwei unvergleichbare Elemente von T_0 sind, so sind $E(a)$, $E(b)$ auch zwei unvergleichbare Elemente.

Kurz: die Zuordnung $a \mapsto E(a)$, ($a \in T_0$) ist eine Ähnlichkeit der Tafel T_0 mit der Tafel $E(T_0)$ aller $E(a)$, ($a \in T_0$); nun ist nach dem Satz 1 $E(T_0)$ normal; dasselbe gilt also von T_0 und also auch von T , womit der Satz 3 bewiesen ist.

9. Dritter Fall. Tafel E_ω ($E =$ beliebige separable geordnete Menge ³⁾).

Sei

(16) M

die geordnete Menge aller geordneten Paare (x, a) mit $0 \leq x \leq 1$ und $a = 2$ oder 3, in der für zwei verschiedene Elemente (x, a) , (x', a') das Symbol $(a, b) < (a', b')$ gleichbedeutend ist entweder mit $x = x'$, $a < a'$ oder mit $x < x'$.

M ist eine geordnete separable Menge (vom Typus 2θ), und jede geordnete separable Menge ist ähnlich zu einer Untermenge von M^{12} .

Wenn daher M die soeben definierte Menge bezeichnet, wird der dritte Fall durch folgenden Satz erledigt werden:

Satz 4. Jede Tafel $T \subseteq M_\omega$ ist normal.

Sei nun für irgendwelche nichtleere wohlgeordnete Teilmenge A von M

$$f(A) = \sup_x,$$

x durchlaufend die Menge der Zahlen ≤ 1 , die in den Elementen der Menge A vorkommen. Dann ist $0 \leq f(A) \leq 1$ und $f(A) \leq f(B)$ für jedes andere Element $B \in M_\omega$ mit $A < B$, und sogar $f(A) < f(B)$, wenn B noch mindestens zwei Elemente mehr als A enthält. Wie früher kann man auch jetzt $\gamma T = \Omega$ voraussetzen; wenn dann $T_2 = \sum_{\alpha} R_{\omega\alpha} T$, ($\alpha < \Omega$), so ist $f(x)$, ($x \in T_2$) eine in T_2 wachsende reelle Funktion; nach Satz 3 ist also T_2 normal; dasselbe gilt von T , weil T_2 eine mit T gleichmächtige Untermenge von T ist.

¹²⁾ D. h. 2θ ist ein universaler, geordneter, separabler Ordnungstypus. Vgl. Georges Kurepa, Sur les relations d'ordre. Bull. International Ac. Yougoslave des Sciences, Zagreb 1939, livre XXXII, S. 66–76 (Satz 9b).

10. *Vierter Fall. Tafel E_w [E = beliebige geordnete mit Suslinscher Eigenschaft behaftete Menge ⁴⁾].*

Satz 5. *Wenn eine geordnete mit Suslinscher Eigenschaft behaftete Menge E so beschaffen ist, daß jede Tafel $T \subseteq E_w$ normal ist, so ist M separabel³⁾ (und daher von einem Ordnungstypus $\leq 2 \vartheta$).*

Man sieht leicht ein, daß es keine Einschränkung ist, die Menge E als lückenlos und mit beiden Endpunkten behaftet vorauszusetzen.

Definieren wir dann die Folge

$$(17) \quad D_0, D_1, \dots, D_\xi, \dots$$

von Systemen der *Stücke*¹³⁾ von E wie folgt:

D_0 ist ein System bestehend aus der Menge E selbst; wenn $0 < \xi$ und wenn die Systeme $D_0, D_1, \dots, D_\alpha, \dots$ ($\alpha < \xi$) so definiert sind, daß in jedem D_α mindestens ein mehrpunktiges Stück von E vorhanden ist, so wird D_ξ folgendermaßen definiert: Ist ξ von 1. Art, so wird jedes mehrpunktige Element von $D_{\xi-1}$ in zwei *disjunkte nichtleere Stücke* zerlegt; die Menge aller dieser Teilstücke sei D_ξ ; wenn ξ von 2. Art ist, so bedeute D_ξ die Menge aller Durchschnitte

$$X_0 X_1 \dots X_\alpha \dots, \quad (\alpha < \xi)$$

mit

$$X_0 \supset X_1 \supset \dots \supset X_\alpha \supset \dots, \quad X_\alpha \in D_\alpha, \quad (\alpha < \xi).$$

Sei γ die erste Zahl, für welche das System D_γ leer ist; wenn dann

$$(18) \quad D$$

das System aller nichtleeren Elemente aller Systeme D_ξ , ($\xi < \gamma$) bezeichnet, so beweist man leicht:

a) D ist eine bezüglich \supset verzweigte Tafel, in der *unvergleichbare* Elemente mit *disjunkten* Mengen zusammenfallen; es ist

$$\gamma D = \gamma, \quad R_\xi D = D_\xi, \quad (\xi < \gamma).$$

b) Die Elemente von D_ξ sind punktfremde Stücke von E .

c) Jede Klasse von 1. Art von D ist höchstens abzählbar.

d) Die Menge der Endpunkte aller Elemente 1. Art von D ist überall-dicht in E .

Um daher die Separabilität von E zu beweisen, genügt es einzusehen, daß

$$(19) \quad \gamma D < \Omega,$$

oder, daß die Relation

$$(20) \quad \gamma D \geq \Omega$$

zu einem Widerspruche führt. Nehmen wir also an, daß $\gamma D \geq \Omega$.

¹³⁾ Jede Teilmenge von E , die mit zwei Punkten von E auch alle dazwischenliegenden Punkte von E enthält, heißt ein Stück von E .

11. Sei

$$(21) \quad \psi D$$

das System aller *mehrpunktigen* Elemente aus D .

a) ψD ist eine Teiltafel von D und wegen (20) $\gamma D = \gamma \psi D$, also nach (20)

$$\gamma \psi D \geq \Omega.$$

b) Jede Klasse von ψD ist höchstens abzählbar [vgl. 10., c)].

12. Sei für jedes $A \in \psi D$

$$(22) \quad \alpha(A) = \inf_x x, \quad (x \in A),$$

d. h. $\alpha(A)$ sei der letzte Punkt von E , so daß kein Punkt von A dem Punkte $\alpha(A)$ vorangeht; wegen der Lückenlosigkeit und Begrenztheit von E ist der Punkt $\alpha(A)$ wohlbestimmt.

a) Zu jedem $x \in E$ gibt es höchstens abzählbar viele $A \in \psi D$ mit $\alpha(A) = x$; sonst gäbe es in jeder Klasse $R_\xi \psi D$, ($\xi < \Omega$) ein bestimmtes A_ξ mit $\alpha(A_\xi) = x$; wenn dann für jede Zahl $\xi < \Omega$ a_ξ irgendeinen Punkt aus $A_\xi - A_{\xi+1}$ bezeichnet, so hätte man die transfinite absteigende Folge $a_0, a_1, \dots, a_\xi, \dots$ verschiedener Punkte aus E , was der Suslinschen Bedingung widerspricht.

b) Die Tafel ψD und die Menge $\alpha(\psi D)$ aller $\alpha(X)$, ($X \in \psi D$), sind gleichmächtig.

Es ist nämlich $p\alpha(\psi D) \leq p\psi D^{10)}$; andererseits nach a) $p\psi D \leq p\omega p\alpha(\psi D)$, daher $p\psi D = p\alpha(\psi D) = p\Omega$.

13. Sei für jedes $A \in \psi D$

$$(23) \quad E(A)$$

die Menge aller $\alpha(X)$ mit $X \in \psi D$, $X \supseteq A$.

a) $E(A)$ ist eine wohlgeordnete Teilmenge von E ; also $E(A) \in E_\omega$.

b) Das System

$$(24) \quad E(\psi D)$$

aller $E(X)$, ($X \in \psi D$) ist eine verzweigte, mit der Menge $\alpha(\psi D)$ [vgl. 12., b)] gleichmächtige Tafel.

Sei nämlich, für $x \in \alpha(\psi D)$, A ein beliebiges Element aus ψD mit $\alpha(A) = x$; wenn dann $\varphi(x) = E(A)$, so ist die Zuordnung $x \leftrightarrow \varphi(x)$ *eindeutig*. Zunächst, wenn $\alpha(A) = \alpha(B)$, $A \neq B$, dann ist für jedes zwischen A und B liegende Element C auch $\alpha(C) = x$, also $E(C) = E(A) = E(B)$; somit ist $\varphi(x)$ eindeutig. Wenn weiter x, x' zwei verschiedene Punkte von $\alpha(\psi D)$ sind, so ist $\varphi(x) \neq \varphi(x')$. Sei nämlich $x = \alpha(A)$, $x' = \alpha(A')$; je nachdem ob $A \subset A'$, $A \supset A'$ oder $A \cap A' = 0$, gehört x nicht zu $\varphi(x')$, oder x' nicht zu $\varphi(x)$, oder es gilt das eine und das andere.

c) $E(\psi D)$ ist gleichmächtig mit ψD . Das folgt aus b) und [12, b)].

14. Der Voraussetzung nach ist $E(\psi D) \subseteq E_w$ eine normale Tafel; sei dann T eine mit $E(\psi D)$ gleichmächtige degenerierte Untertafel von $E(\psi D)$; sei, für jedes $t \in T$, t_0 ein bestimmtes Element von ψD mit $E(t_0) = t$; sei ferner T_0 die Menge aller t_0 , ($t \in T$).

a) T_0 wäre eine abzählbare degenerierte Tafel $\subseteq \psi D$. Wäre T_0 nicht degeneriert, so gäbe es also drei Elemente A, B, C aus T_0 mit $C \supset A$, $C \supset B$, $AB = 0$ (A, B, C sind doch verschiedene Stücke der Menge E), so hätte man für die zugehörigen Elemente von T $E(C) \leq E(A)$, $E(C) \leq E(B)$, $E(A) \neq E(B)$, und weder $E(A) < E(B)$ noch $E(B) < E(A)$, was der Entartung von T widerspricht.

b) Die Klasse $R_0 T_0$ enthielte eine abzählbare Familie punktfremder, mehrpunktiger Stücke von E .

Wird dann für ein $A \in R_0 T_0$ mit A^0 die Menge aller $X \in T_0$ mit $A \supseteq X$ bezeichnet, so hätten wir folgendes:

α) A^0 ist eine absteigende Folge von Elementen aus ψD ; daher ist A^0 höchstens abzählbar;

β) Jedes Element von T_0 ist in einer und nur einer Familie A^0 enthalten.

Aus α) und β) folgt $p T_0 \leq p R_0 T_0 p \omega^{10}$, was mit der Voraussetzung (20) oder $p T_0 = p \Omega$ die unmögliche Tatsache b) ergäbe. Daher ist die Annahme (20) falsch, womit die Separabilität der Menge E bewiesen ist.

15. Weil jede geordnete separable Menge die Suslinsche Eigenschaft besitzt, können die Sätze 4 und 5 folgendermaßen zusammengesetzt werden:

Die notwendige und hinreichende Bedingung, damit eine geordnete mit Suslinscher Eigenschaft behaftete Menge E eine höchstens abzählbare, überalldichte Teilmenge enthält, ist die, daß jede Teiltafel aus E_w [vgl. (5)] normal ist.

(Eingegangen am 22. 6. 1941.)

Zur Frage des Euklidischen Algorithmus in quadratischen Zahlkörpern.

Von

L. Rédei in Szeged (Ungarn).

Es sei $K = R(\sqrt{m})$ ein beliebiger quadratischer Zahlkörper über dem Körper R der rationalen Zahlen, $m (\neq 0, 1)$ eine quadratfreie ganz rationale Zahl. Es werde wie bei allen Autoren folgende Definition von Dickson [4]¹⁾ zugrunde gelegt.

a. In K existiert der Euklidische Algorithmus (kurz E.A.), wenn es zu allen Paaren α, μ ($\mu \neq 0$) von ganzen Zahlen in K eine dritte ganze Zahl β in K mit

$$(1) \quad |N(\alpha - \mu\beta)| < |N\mu|$$

gibt, wobei N die Norm in K bedeutet.

Bisher weiß man über den E.A. folgendes. Er existiert für $m = -11, -7, -3, -2, -1, 2, 3, 5, 6, 7, 11, 13, 17, 19, 21, 29, 33, 37, 41, 57$. Außerdem kann er nur für endlich viele weitere m existieren, und zwar nur dann, wenn $m = 61, 109$ oder eine positive Primzahl $\equiv 1 \pmod{8}$ ist, insbesondere aber von den letzteren-bis zur Schranke $m \leq 3001$ nur für 73, 89, 97, 113, 137, 193, 241, 313, 337, 457, 601²⁾.

Für negative m ist das Resultat von Dickson [4] angegeben. Für positive m haben die Existenz des E.A. in den angegebenen Fällen Dickson [4], Perron [12], Oppenheim [11], Berg [2] und Hofreiter [9] bewiesen, gewisse Fälle hiervon sind von Remak [15], Behrbohm und Rédei [1] behandelt. Unvergleichbar mehr Mühe haben die Nichtexistenzbeweise gekostet. Und zwar haben Behrbohm und Rédei [1] bewiesen, daß der E.A. nur in den folgenden vier Fällen existieren kann:

- I. $m = p \equiv 1 \pmod{4}$,
- II. $m = pq$, $p \equiv q \equiv 3 \pmod{4}$,
- III. $m = 2$ oder $m = 2p$, $p \equiv 3 \pmod{4}$,
- IV. $m = p \equiv 3 \pmod{4}$,

¹⁾ Zahlen in eckigen Klammern beziehen sich auf das am Ende der Arbeit angeführte Literaturverzeichnis.

²⁾ Für diese m und $m = 61, 109$ ist die Klassenzahl in K gleich 1. Eine obere Schranke für die m mit E.A. ist nicht bekannt.

wobei p und q positive Primzahlen sind. Für die Fälle III und IV hat aber Berg [2] schon früher bewiesen, daß sie keine weiteren m mit E.A. enthalten. Andere Beweise sind von Fox Keston [7] und von Behrbohm und Rédei [1]. Letztere haben auch für den Teil $p \equiv 5 \pmod{24}$ von I und für den Teil $m \equiv 5 \pmod{8}$, $m \neq 3p$ von II, dann aber Hofreiter [9] auch für $m \equiv 5 \pmod{8}$, $m \neq 3p$ die ähnlichen Resultate bekommen. Auf analytischem Wege haben Erdős und Ko [6] bewiesen, daß der E.A. im Fall I nur für endlich viele $p \equiv 5 \pmod{24}$ existieren kann. Für Fall II mit $m \equiv 1 \pmod{8}$ hat Heilbronn [8] das Analoge geleistet, wobei von ihm auch der Fall von Erdős und Ko mitbehandelt wurde. Brauer [3] verfeinerte das Resultat von Erdős und Ko für $p \equiv 13 \pmod{24}$ auf elementarem, aber kompliziertem Wege durch die Angabe der Schranke $p < 3300000$ für diese p mit E.A., weiter ergab eine mit Hilfe von zwei Mitarbeitern (Mutter und Frau) durchgeführte Nachprüfung (s. unten) der einzelnen Fälle die Schranke $p \leq 109$. Heilbronn's Resultat für Fall II übertrafen Schuster [16] und Rédei [13], die im restlichen Fall $m \equiv 1 \pmod{8}$ für $m = 3p$ und $m \neq 1 \pmod{24}$ bzw. für $m \equiv 1 \pmod{24}$ bewiesen haben, daß es dann kein m mit E.A. außer den oben angegebenen gibt: Der Beweis in der letzteren Arbeit erstreckt sich übrigens auf alle m im Fall II mit der einzigen Ausnahme $m = 3p$ und ist elementar bis auf die Anwendung eines analytisch gewonnenen Satzes von Hasse³⁾ in der Theorie der Verteilung der quadratischen Reste. Bei Rédei [14] wurde der Beweis völlig elementar geführt.

Wir wiederholen, daß außer Fall I mit $p \equiv 1 \pmod{8}$ oder $p = 61, 109$ kein Problem mehr übrig ist und alles elementar gewonnen wurde bis auf den diesbezüglichen Teil des Endlichkeitssatzes von Erdős und Ko. Das läßt vermuten, daß die volle Lösung (wenn überhaupt) elementar entstehen wird. (In vorliegender Arbeit schließen wir auch nur elementar.)

Bemerken wir noch, daß im Fall I alle bisherigen Nichtexistenzbeweise (auch der Endlichkeitssatz von Erdős und Ko) aus folgendem Lemma (Behrbohm und Rédei [1], S. 198, 1., s. unten) entstanden sind:

b. Gibt es im Fall I eine Zerlegung $p = ab + cd$ (a, b, c, d positiv ganz) mit $(a, b) = (c, d) = 1$, $\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right) = \left(\frac{c}{p}\right) = -1$ (also auch $\left(\frac{d}{p}\right) = -1$), so gilt der E.A. nicht.

Insbesondere wurde das Intervall $109 < p < 3300000$ [$p \equiv 13 \pmod{24}$] bei Brauer [3] mit dieser Hilfe überprüft. Ebenfalls durch b ergab sich, daß der E.A. im Intervall $73 \leq p \leq 3001$ [$p \equiv 1 \pmod{8}$] höchstens für die oben

³⁾ H. Hasse, Abstrakte Begründung der komplexen Multiplikation und Riemannsche Vermutung in Funktionenkörpern. Abh. Math. Sem. Hamburg 40 (1934), S. 347.

angeführten elf Fälle 73, ..., 601 existieren kann⁴⁾. Dagegen sind diese p und $p = 61, 109$ lauter Fälle, in denen b keinen Erfolg bringt.

Da es uns unmöglich schien, daß alle diese 13 Fälle solche mit E.A. sind, suchten wir nach einem anderen (von b verschiedenen) Weg, um auf Nichtexistenz des E.A. zu schließen. So haben wir bekommen, daß $p \equiv 17 \pmod{24}$ ($p > 41$) und $p = 61, 109$ ⁵⁾ ohne E.A. sind. Dagegen gilt der E.A. für $p = 73$. (Der Beweis hierfür ist keine Neuerung in der Methode.) Insbesondere folgt hieraus, daß das größte m mit E.A. eine Primzahl $[\equiv 1 \pmod{24}, \geq 73]$ ist. (Bisher war bis zu der zusammengesetzten Zahl 57 ein E.A. erwiesen.) Zu untersuchen bleibt also nur noch der Fall I mit $p \equiv 1 \pmod{24}$, $p \geq 97$, insbesondere bis zur Schranke $p \leq 3001$ nur $p = 97, 193, 241, 313, 337, 457, 601$.

Es ist ein enger Zusammenhang mit der Grundeinheit $\varepsilon = t + u\sqrt{p}$ (> 1 ; t, u rational) in den Vordergrund getreten, den wir aber nicht fein genug verfolgen konnten⁶⁾. Insbesondere scheint ein großes ε das Problem zu erschweren. Für festes u ließe sich leicht eine obere Schranke für die $p [\equiv 1 \pmod{8}]$ mit E.A. gewinnen. Das haben wir (zu weiteren Zwecken) nur für $u = 1$ ausgeführt mit dem Resultat, daß dann $p (> 17)$ ohne E.A. ist (vgl. 5.).

Um alles in klare Beleuchtung zu setzen, wollen wir zuerst auseinander setzen, welchen Ausgangspunkt Perron [12] und nach ihm auch alle späteren Autoren gewählt haben. Er dividierte (1) durch $|N\mu|$, wodurch a folgende Form erhält:

a^I. In K existiert der E.A. dann und nur dann, wenn es zu jeder (gebrochenen) Körperzahl γ eine ganze Körperzahl β mit

$$(2) \quad |N(\gamma - \beta)| < 1$$

gibt.

Durch Einführung einer Basisdarstellung entsteht hieraus Perrons Kriterium:

a^{II}. In K existiert der E.A. dann und nur dann, wenn es zu jedem rationalen Zahlenpaar x, y ein ganz rationales Zahlenpaar a, b mit

$$(3) \quad \left| \left(x - \frac{a}{2} \right)^2 - m \left(y - \frac{b}{2} \right)^2 \right| < 1 \begin{cases} a \equiv b \pmod{2} & \text{für } m \equiv 1 \pmod{4} \\ a \equiv b \equiv 1 \pmod{2} & \text{für } m \not\equiv 1 \pmod{4} \end{cases}$$

gibt.

⁴⁾ Über 3001 liegt bisher keine ähnliche Nachprüfung vor.

⁵⁾ Diese Resultate für $p = 61, 109$ waren mir schon vor einem Jahre bekannt, durch die eben die Lücken in Brauers Arbeit ausgefüllt werden.

⁶⁾ Eine Erwähnung eines solchen Zusammenhangs kommt bisher nur bei Remak [15] vor.

Nennt man die $(\frac{a}{2}, \frac{b}{2})$ Gitterpunkte, bezeichnet ferner H das (offene)

Hyperbelkreuz

$$(4) \quad |x^2 - my^2| < 1$$

mit den Halbachsen 1, $\frac{1}{\sqrt{m}}$ und H_G das Hyperbelkreuz, das den Gitterpunkt G zum Mittelpunkt hat und durch Verschiebung aus H entsteht, so lautet a^{II} geometrisch:

a^{III} . In K existiert der E. A. dann und nur dann, wenn die H_G alle rationalen Punkte⁷⁾ der Ebene überdecken. (Etwas anders bei Remak [15].)

Alle bisherigen Existenzbeweise lassen sich am bequemsten durch a^{III} gewinnen, und so werden wir auch den Fall $p = 73$ behandeln.

Die bisherigen Nichtexistenzbeweise sind aber (mehr oder weniger ausdrücklich) so aus a^{II} entstanden, daß man die Unmöglichkeit von (3) für spezielle x, y nachgewiesen hat⁸⁾, und zwar konnte man dabei mit $x = 0$, $y = \frac{z}{m}$ (z ganz rational) auskommen⁹⁾, nur im einzigen Fall $m = 14$ wäre das ohne Erfolg, dann hilft aber $x = y = \frac{1}{2}$.

Wir wollen hier unmittelbar auf Definition a zurückgreifend folgenden Weg einschlagen. Bei festem μ soll $\bar{\mu}$ immer eine beliebige ganze Zahl in K mit $|N\bar{\mu}| < |N\mu|$ bedeuten. Dann heißt (1), daß eine Gleichung $\alpha - \beta\mu = \bar{\mu}$, d. h. eine Kongruenz $\alpha \equiv \bar{\mu} \pmod{\mu}$ besteht. Also läßt sich a so aussagen:

a^{IV} . In K existiert der E. A. dann und nur dann, wenn für jede ganze Zahl $\mu (\neq 0)$ in K die Restklassen $\bar{\mu} \pmod{\mu}$ alle ganzen Zahlen von K enthalten. Je nachdem diese Bedingung für μ erfüllt ist oder nicht, nennen wir es einen euklidischen bzw. nichteuklidischen Modul. Das Vorhandensein eines nichteuklidischen Moduls ist also charakteristisch für das Fehlen des E. A.

Da die Einheiten in K keine nichteuklidischen Moduln sind, wollen wir fortan annehmen, daß μ weder Null noch eine Einheit ist. Offenbar gilt folgendes:

c. Ist μ ein nichteuklidischer Modul, so gilt dasselbe für das Konjugierte und für alle Assoziierten und Vielfachen von ihm.

Es ist leicht zu sehen (s. unten Näheres darüber), daß das Einsetzen der obenerwähnten Zahlenpaare $x = 0, y = \frac{z}{m}$ bzw. $x = y = \frac{1}{2}$ in a^{II} darauf hinauskommt, daß man für μ in a^{IV} einen passenden Teiler von \sqrt{m} oder

⁷⁾ D. h. Punkte mit rationalen Koordinaten.

⁸⁾ Nur die Einschränkung auf die Fälle I–IV ist anders, nämlich daraus entstanden, daß die Klassenzahl in einem K mit E. A. notwendig gleich 1 ist (vgl. Behrbohm u. Rédei [1], S. 193).

⁹⁾ Dabei war im allgemeinen $(z, m) = 1$ angenommen, nur für Fall II $(z, m) = (z, pq) = p$ oder q . Es ist interessant, daß im Fall II die scheinbar stärkere Annahme $(z, m) = 1$ bei Heilbronn [8] (sogar auf analytischem Wege) zu keinen Vollresultaten führen konnte.

$\mu = 2$ wählt¹⁰⁾. Insbesondere bedeutet das für Fall I, daß man auf b geschlossen hat. Nach dem Obigen muß also auch nach anderen μ gegriffen werden, will man in den angeführten Fällen die Nichtexistenz des E.A. ausweisen. Aus einfachen Gründen, auf die wir später noch hinweisen, werden wir in dieser Arbeit $\mu = \varepsilon + 1$ wählen. Damit will nicht gesagt werden, daß weitere μ keineswegs zu empfehlen wären; insbesondere ließ sich aber für $p = 61$ beweisen, daß dann die Nichtexistenz des E.A. nur mit Hilfe von $\mu = \varepsilon + 1$ zu bestätigen ist (d. h. nach c treten noch unendlich viele μ hinzu, die trivial auch dasselbe leisten).

Geometrisch läßt sich das Gesagte folgendermaßen interpretieren. Wir ordnen die Zahlen $x + y\sqrt{m}$ (x, y rational) und die Punkte (x, y) einander zu. Für eine beliebige Zahl γ in K nennen wir alle Zahlen in K äquivalent, die $\equiv \pm \gamma$ oder $\pm \gamma'$ (mod 1) sind — wobei der Strich das Konjugierte in K bedeutet — und ebenso die entsprechenden Punkte. Für die letzteren bildet das Rechteck zwischen den Geraden $y = 0, y = \frac{1}{2}, x = 0, x = \frac{1}{2}$ (Fälle I, II) bzw. das Quadrat zwischen den Geraden $y = 0, y = \frac{1}{2}, x = 0, x = \frac{1}{2}$ (Fälle III, IV) einen Fundamentalbereich T im üblichen Sinne. Unlösbarkeit von (2) bedeutet offenbar, daß der γ zugeordnete Punkt im Sinne von $\mathfrak{a}^{\text{III}}$ unüberdeckbar ist. Dabei verhalten sich äquivalente Zahlen bzw. Punkte gleich. Also gilt $\mathfrak{a}^{\text{III}}$ auch für T statt der Ebene. Nach den obigen lassen sich alle bisherigen Nichtexistenzfälle geometrisch so ausweisen, daß man einen rationalen Punkt auf der Grenze von T und zwar im allgemeinen einen auf der y -Achse, im Fall $m = 14$ den Eckpunkt $(\frac{1}{2}, \frac{1}{2})$ angibt, den kein H_0 überdeckt (bzw. man beweist die Existenz eines von ihnen). In vorliegender Arbeit suchen wir dagegen nach unüberdeckbaren rationalen Punkten im Innern von T ¹¹⁾.

Zuerst wollen wir die Existenz des E.A. für $p = 73$ beweisen. Bei den darauffolgenden Nichtexistenzbeweisen werden wir (durch die Annahme $\mu = \varepsilon + 1$ in \mathfrak{a}^{IV}) ein Lemma g erhalten, das wir dann auf $p = 61, 109$ und $p \equiv 17 \pmod{24}$ ($p > 41$) in dieser Reihenfolge anwenden. Dabei wird g nur sehr wenig ausgenutzt, und das läßt weitere Anwendungsmöglichkeiten hoffen. Um den Weg für weitere Untersuchungen zu ebnen, entwickeln wir manches etwas allgemeiner, als das unbedingt nötig ist.

* * *

¹⁰⁾ Mit anderen Worten suchte man bisher nach einem nichteuklidischen Modul nur unter den Zahlen μ mit $\mu \mid \sqrt{m}$ und $\mu = 2$. Diese Wahl war allerdings sehr glücklich, wenn sie im Fall I auch nicht alles leisten konnte (s. unten).

¹¹⁾ Es hat sich sogar herausgestellt, daß für $p = 61$ nur ein einziger unüberdeckbarer rationaler Punkt im Fundamentalbereich T vorhanden ist, nämlich der innere Punkt $(\frac{1}{78}, \frac{17}{78})$, ein klarer Beweis, daß man auf der alten Spur nicht alles erhalten kann.

1. Wir beweisen, daß der E.A. in $K = R(\sqrt{73})$ existiert. Es genügt nach a^{III} zu zeigen, daß die $H_{G_i} = H_i$ mit den Gitterpunkten $G_i = (\frac{a_i}{2}, \frac{b_i}{2})$

$i =$	1	2	3	4	5	6	7	8	9
$a_i =$	0	-2	23	-54	57	-20	14	5	-4
$b_i =$	0	0	3	-6	7	-2	2	1	0

alle rationalen Punkte des durch die Geraden $x = 0$, $x = \frac{1}{2}$, $y = 0$, $y = x(x - \frac{1}{2}) + \frac{1}{4}$ begrenzten Trapezes T_1 enthalten, wobei $x = \frac{1}{\sqrt{73}}$ ist¹²⁾.

H_i besteht aus den Punkten (x, y) mit

$$\left| \left(x - \frac{a_i}{2} \right)^2 - 73 \left(y - \frac{b_i}{2} \right)^2 \right| < 1.$$

Setzt man also für $0 \leq x \leq \frac{1}{2}$ ¹³⁾

$$f_{i,\sigma}(x) = \frac{1}{2} (b_i + \sigma_i x \sqrt{(2x - a_i)^2 + 4\sigma\sigma_i}) \quad (i = 1, \sigma = 1; i = 2, 3, \dots, 9, \sigma = \pm 1),$$

wobei $\sigma_i = 1$ für $a_i \leq 0$ und $\sigma_i = -1$ für $a_i > 0$ ist, so ist die Punktmenge (x, y) definiert durch $0 \leq x \leq \frac{1}{2}$ und durch je eine der Ungleichungen

$$0 \leq y < f_{1,1}(x) \quad \text{und} \quad f_{i,-1}(x) < y < f_{i,1}(x) \quad (i = 2, 3, \dots, 9)$$

in H_1 bzw. H_i enthalten. Es genügt also zu zeigen, daß die Funktionen

$$g_i(x) = f_{i,1}(x) - f_{i+1,-1}(x) \quad (i = 1, 2, \dots, 8), \quad g_9(x) = f_{9,1}(x) - x(x - \frac{1}{2}) - \frac{1}{4}$$

im Intervall $0 \leq x \leq \frac{1}{2}$ positiv sind bis auf $g_1(\frac{1}{2}) = 0$ mit irrationalem $f_{1,1}(\frac{1}{2}) (= f_{2,-1}(\frac{1}{2}) = \frac{x}{2} \sqrt{5})$. [D.h. es bleibt nur der irrationale Punkt

$(\frac{1}{2}, \frac{x}{2} \sqrt{5})$ von T_1 unüberdeckt.] Das folgt daraus, daß für die g_i und die Ableitungen g'_i folgendes gilt: $g_1(\frac{1}{2}) = 0$, $g'_1(x) < 0$ ($0 \leq x \leq \frac{1}{2}$); $g_i(\frac{1}{2}) > 0$, $g'_i(x) < 0$ ($0 \leq x \leq \frac{1}{2}$) für $i = 2, 4, 5, 8$; $g_i(0) > 0$, $g'_i(x) > 0$ ($0 \leq x \leq \frac{1}{2}$) für $i = 3, 6, 7$; $g_9(\frac{1}{2}) > 0$, $g'_9(x) \leq 0$ ($0 \leq x \leq \frac{1}{2}$), $g'_9(x) \geq 0$ ($\frac{1}{2} \leq x \leq \frac{1}{2}$). Damit ist die Behauptung bewiesen.

2. Jetzt wenden wir uns zu den Nichtexistenzbeweisen. Größerer Klarheit halber halten wir am Anfang alle Fälle I bis IV vor Augen. Bei festem μ liegen immer unendlich viele $\bar{\mu}$ vor, da mit einem $\bar{\mu}$ gleichzeitig alle assoziierten auftreten. Wählt man aber aus ihnen ein volles System nichtassoziierter Elemente $\bar{\mu}_1, \bar{\mu}_2, \dots, \bar{\mu}_M$ heraus (bei festem K ist M nur von μ abhängig),

¹²⁾ T_1 ist ebenso wie T ein Fundamentalbereich.

¹³⁾ Die Lücken des von hier an kurz gefaßten Beweises sind nach Behrbohm u. Rédei [1], S. 196–197 leicht auszufüllen. Hier auf S. 197, Zeile 7 von oben, ist

$\frac{1}{x_{i+1}} \geq -\frac{1}{x_i}$ statt $-\frac{1}{x_{i+1}} \geq \frac{1}{x_i}$ zu lesen.

so sind alle $\bar{\mu}$ durch $\eta \bar{\mu}_i$ (η Einheit in K) angegeben. Deshalb definieren wir: Die kleinste natürliche Zahl $\{\mu\}$ mit $\mu | \varepsilon^{(\mu)} + \sigma$ ($\sigma = \pm 1$) heißt der Index von μ . Offenbar gilt dann:

d. Alle Restklassen $\bar{\mu} \pmod{\mu}$ kommen unter den Restklassen $\pm \varepsilon^s \bar{\mu}_j$ ($\pmod{\mu}$) ($i = 0, 1, \dots, \{\mu\} - 1$; $j = 1, 2, \dots, M$) vor.

Will man also einen nichteuklidischen Modul μ finden, so scheint (ein kleiner Index $\{\mu\}$, insbesondere) $\{\mu\} = 1$ von Vorteil zu sein. Noch günstiger kann aber ein solches μ sein, wofür die Restklassen leicht zu handhaben sind, und das trifft vor allem für die Teiler von \sqrt{m} zu. (Dann ist nämlich eine ganze Zahl in K einfach ihrem rationalen Teil kongruent $\pmod{\mu}$). Die größte Leistung erwartet man von einem μ , wofür beides gleichzeitig besteht: $\mu | \sqrt{m}$ und $\{\mu\} = 1$.

Wir wollen diese μ bestimmen, wobei wir nach ε Assoziierte außer acht lassen. Auch nehmen wir an, daß die Klassenzahl gleich 1 ist, da es sonst kein Problem gibt. Dementsprechend gebrauchen wir den Idealbegriff gar nicht und sprechen z. B. über Primzahlen in K . Wie auch früher setzen wir

$$(5) \quad \varepsilon = t + u \sqrt{m} \quad (t, u \text{ rational}).$$

Die gesuchten μ sind die von einer Einheit verschiedenen Teiler von $\delta = (\varepsilon + \sigma, \sqrt{m})$, wobei wieder $\sigma = \pm 1$ ist. Im Fall I ist

$$(6) \quad t^2 - pu^2 = -1,$$

woraus $t \equiv \pm 1 \pmod{p}$, $\delta = 1$ folgt. In diesem Fall gibt es also kein gewünschtes μ . (Daran liegt, daß Fall I schwerer ist als die übrigen.) In den Fällen II bis IV ist

$$t^2 - pu^2 = 1,$$

also gilt für jeden Primteiler π von \sqrt{m} $t \equiv \pm 1 \pmod{\pi}$. Es ist somit jedes $\mu = \pi$ eine gewünschte Zahl, und das gleiche leuchtet im Fall III auch über $\mu = \sqrt{2p}$ ein. Damit sind die gewünschten μ erschöpft¹⁴⁾.

Nun entstehen in der Tat alle Nichtexistenzbeweise in den Fällen II bis IV, ausgenommen $m = 14$, auf der Grundlage, daß man \mathfrak{a}^{IV} im Fall II auf $\mu = (p, \sqrt{pq})$ und $\mu = (q, \sqrt{pq})$, im Fall III auf $\mu = \sqrt{2p}$, im Fall IV auf $\mu = \sqrt{p}$ anwendet¹⁵⁾, d. h. immer auf die „größten“ Teiler von \sqrt{m} vom

¹⁴⁾ Da im Fall II $[\sqrt{pq}] = 2$ ist, wie man das leicht sieht.

¹⁵⁾ Dies geschah bisher nur auf dem in der Einleitung geschilderten „Umwege“ mittels des Perronschen Kriteriums \mathfrak{a}^{II} . In einer methodischen Ausarbeitung der Frage des E.A. ist der obige auch mehr ökonomische Weg durch \mathfrak{a}^{IV} vorzuziehen. Heilbronn [8], S. 522 allein hat bisher einen ähnlichen Weg eingeschlagen (für die Fälle I, II), nahm aber auch für Fall II $\mu = \sqrt{m} = \sqrt{pq}$ [vgl. ¹⁴⁾] und betrachtete nur die primen Restklassen (durch obige μ zieht man eben die nichtprimen heran), vgl. ⁹⁾.

Index 1. Der Ausnahmefall $m = 14$ läßt sich ähnlich durch $\mu = 2$ erledigen, wofür übrigens auch $\{\mu\} = 1$ ist, nur $\mu|\sqrt{m}$ gilt nicht.

Nach dieser Abschweifung wenden wir uns wieder dem Fall I zu. Da es jetzt kein μ mit $\mu|\sqrt{p}$, $\{\mu\} = 1$ gibt, so denkt man in erster Reihe an ein μ , das wenigstens der einen dieser Bedingungen genügt. Wählt man $\mu|\sqrt{p}$, d. h. $\mu = \sqrt{p}$, wofür wegen $s^2 \equiv t^2 \equiv -1 \pmod{\sqrt{p}}$ $\{\mu\} = 2$ ist, so entsteht aus a^{IV} , wie schon erwähnt, eben Lemma b^{16}). Deshalb werden wir im Fall I ein μ mit $\{\mu\} = 1$ wählen¹⁷⁾.

3. Bevor wir dies tun, betrachten wir vorläufig ein beliebiges μ und lassen wieder alle Fälle I bis IV zu. Es bedeute $||\omega||$ (ω in K) das größere von $|\omega|$ und $|\omega'|$. Wir beweisen folgendes:

e. Bei einem beliebigen μ gibt es in jeder Restklasse $\bar{\mu} \pmod{\mu}$ ein solches $\bar{\mu}$, wofür $||\bar{\mu}|| \leq \sqrt{\varepsilon^{(u)}}|N\bar{\mu}|$, um so mehr also $||\bar{\mu}|| < \sqrt{\varepsilon^{(u)}}|N\bar{\mu}|$ gilt.

Setzen wir nämlich $\eta = \varepsilon^{(u)}$. Dann ist $\mu|\eta + \sigma$ ($\sigma = \pm 1$). Betrachten wir eine beliebige Restklasse $\bar{\mu} \pmod{\mu}$ und nehmen gleich an, daß $||\bar{\mu}||$ möglichst klein ist. Wegen $-\sigma\eta\bar{\mu} = -\sigma(\eta + \sigma)\bar{\mu} + \bar{\mu} \equiv \bar{\mu} \pmod{\mu}$ und $-\sigma\eta^{-1}\bar{\mu} = -(\sigma + \eta)\eta^{-1}\bar{\mu} + \bar{\mu} \equiv \bar{\mu} \pmod{\mu}$ ist dann

$$(7) \quad ||\eta\bar{\mu}|| \geq ||\bar{\mu}||,$$

$$(8) \quad ||\eta^{-1}\bar{\mu}|| \geq ||\bar{\mu}||.$$

Ist nun erstens $||\bar{\mu}|| = |\bar{\mu}|$, so folgt wegen $|\eta^{-1}\bar{\mu}| < |\bar{\mu}|$ und (8) $|(\eta^{-1}\bar{\mu})'| \geq |\bar{\mu}|$, d. h. wegen $\eta\eta' = \pm 1$: $|\eta\bar{\mu}'| \geq |\bar{\mu}|$, $|\eta N\bar{\mu}| \geq |\bar{\mu}|^2 = ||\bar{\mu}||^2$. Ist zweitens $||\bar{\mu}|| = |\bar{\mu}'|$, so folgt wegen $|\eta^{-1}\bar{\mu}'| < |\bar{\mu}'|$, d. h. wegen $|(\eta\bar{\mu})'| < |\bar{\mu}'|$ und (7) $|\eta\bar{\mu}| \geq |\bar{\mu}'|$, $|\eta N\bar{\mu}| \geq |\bar{\mu}'|^2 = ||\bar{\mu}||^2$. Damit haben wir e bewiesen.

Da es nur endlich viele ganze Zahlen ω in K mit beschränktem $||\omega||$ gibt, die man leicht angeben kann, so ist mit Hilfe von e eine prinzipiell leichte Aufgabe, bei festem μ alle Restklassen $\bar{\mu} \pmod{\mu}$ aufzustellen und also zu entscheiden, ob μ ein nichteuklidischer Modul ist. Schwer ist aber im allgemeinen die wirkliche Ausführung.

4. Nachher wollen wir nur noch den Fall I betrachten. Wie in 2. gesagt, nehmen wir $\{\mu\} = 1$, d. h. $\mu|\varepsilon \pm 1$ an. Da $\varepsilon - 1$ zu $\varepsilon' + 1$ assoziiert ist,

¹⁶⁾ Vgl. 15). Näher führen wir das nicht aus.

¹⁷⁾ Im Fall I sind alle Zahlen μ mit $\{\mu\} = 2$ Teiler von $\left(\frac{\varepsilon^2 + 1}{\varepsilon}\right) = 2u\sqrt{p}$ oder $\left(\frac{\varepsilon^2 - 1}{\varepsilon}\right) = 2t$. Die Aussichten mit einem solchen μ haben wir nicht genügend untersucht. Da auch schon $\{\sqrt{p}\} = 2$ ist und nach e von $\mu = 2u\sqrt{p}$ im allgemeinen mehr als von $\mu = \sqrt{p}$ zu erwarten ist, würde man — um die großen Vorteile des $\mu = \sqrt{p}$ möglichst beizubehalten — vielleicht die Wahl $\mu|2u\sqrt{p}$ bevorzugen. Hier bemerken wir, daß die Annahme $\mu = 2\sqrt{p}$ in a^{IV} ebenfalls nur auf b führt.

verlieren wir nach ϵ an Allgemeinheit nichts, wenn wir direkt $\mu = \epsilon + 1$ setzen. Leicht gewinnen wir aus ϵ folgendes:

1. Im Fall I gibt es für $\mu = \epsilon + 1$ in jeder Restklasse $\bar{\mu} \pmod{\epsilon + 1}$ auch ein solches $\bar{\mu}$, wofür

$$(9) \quad \bar{\mu} = \frac{x + y\sqrt{p}}{2} \quad (x, y \text{ ganz rational, } 2 \mid x + y)^*$$

$$(10) \quad |x^2 - py^2| < 8t,$$

$$(11) \quad |x| \leq 2t + 1, \quad |y| \leq 2u$$

gilt.

Es ist nämlich

$$(12) \quad N(\epsilon + 1) = 2t.$$

Wenn wir also μ in der Form (9) annehmen, so gilt wegen $|N\bar{\mu}| < |N(\epsilon + 1)|$ auch (10). Nach ϵ läßt sich $\bar{\mu}$ wegen (12) durch $\|\bar{\mu}\| < \sqrt{2t\epsilon}$ einschränken. Da

$$(13) \quad 2t < \epsilon < 2u\sqrt{p}$$

ist, so folgt

$$(14) \quad \|\bar{\mu}\| < \epsilon.$$

Nun zeigen wir, daß (11) eine Folgerung aus (9), (10) und (14) ist, wobei es genügt, den Fall $x \geq 0, y \geq 0$ zu betrachten. Nach (9) und (14) ist dann $x < 2\epsilon - y\sqrt{p}$. Durch Quadrieren folgt $4y\epsilon\sqrt{p} < 4\epsilon^2 - (x^2 - py^2)$, also nach (10) $y\epsilon\sqrt{p} < \epsilon^2 + 2t$ und nach (13) $y\epsilon\sqrt{p} < 2u\epsilon\sqrt{p} + \epsilon$, $y < 2u + \frac{1}{\sqrt{p}}$. Da y ganz ist, ist die zweite Ungleichung in (11) richtig. Wenn die erste falsch wäre, so würde daraus $x^2 - py^2 \geq (2t + 2)^2 - p(2u)^2 = 8t$ folgen. Das ist nach (11) unmöglich, und das beweist f.

5. Ist die Anzahl der mit (9), (10), (11) angegebenen $\bar{\mu}$ kleiner als $2t$, so folgt nach (12) und a^{IV} , daß der E. A. im betreffenden K nicht existiert. Auf diesem Weg ist leicht eine von u abhängende obere Schranke für die p mit E. A. zu ermitteln, die aber zu groß, nämlich von der Größenordnung $(\log u)^2$ ist¹⁸⁾. Hier wollen wir nur beweisen, daß es für $u = 1$ außer 17 kein $p \equiv 1 \pmod{8}$ mit E. A. gibt¹⁹⁾. Dies wird uns ermöglichen, f (in 6.) etwas zu verschärfen.

¹⁸⁾ Hätte man die „etwas“ bessere Abschätzung $p < a(\log u)^{2-b}$ (a, \dots, d sind absolute positive Konstanten), so wäre $\log \epsilon > (\log u >) \epsilon p^{\frac{1}{2}+d}$, und das würde wegen der bekannten Abschätzung für $\log \epsilon$ eine absolute Schranke für die p mit E. A. ergeben.

¹⁹⁾ Die ähnliche Aussage (wieder mit Ausnahme der bekannten Fälle mit E. A.) gilt auch für den ganzen Fall I mit $u \leq 1$, da dann nur p mit $p \equiv 5 \pmod{8}$, $p \neq 61, 109$ hinzutreten können.

Es handelt sich nämlich um die $p = a^2 + 1$ mit $4|a$, $\varepsilon = a + \sqrt{p}$, $t = a$. Dann werden die $\bar{\mu}$ in **f** durch (9) mit

$$y = 0, \quad |x| < \sqrt{8a} \quad (2|x|)$$

$$y = \pm 1, \quad |x^2 - a^2 - 1| < 8a \quad (2 \nmid x)$$

$$y = \pm 2, \quad |x^2 - 4a^2 - 4| < 8a \quad (2|x|)$$

geliefert. Für $a > 16$ sind die Lösungen dieser Ungleichungen der Reihe nach die geraden Zahlen zwischen $-\sqrt{8a}$ und $\sqrt{8a}$, dann $\pm(a \pm 1)$, $\pm(a \pm 3)$, endlich $\pm 2a$. Die Anzahl all dieser μ ist $< \sqrt{8a} + 21 < 2a = 2t$ woraus nach obiger Bemerkung die Nichtexistenz des E.A. folgt. Dann ist die Behauptung richtig, da es mit $a \leq 16$ außer 17 und 257 kein p gibt und $p = 257$ ohne E.A. ist.

6. Nach 5. mit der Verschärfung¹⁹⁾ läßt sich fortan $u \geq \frac{3}{2}$ annehmen, und dann gilt folgendes:

f. In **f** darf man (11) durch

$$(11') \quad |x| < 2t, \quad |y| < 2u$$

ersetzen.

Ist nämlich $|y| = 2u$, so ist wegen (9) $|x| \neq 2t \pm 1$. Auch $|x| = 2t$ läßt sich ausschließen, da dann nach (12) $\bar{\mu} = \pm t \pm u\sqrt{p} \equiv \pm t \pm (t+1) \equiv \pm 1 \pmod{\varepsilon+1}$ ist, und diese Restklassen werden auch schon durch $\bar{\mu} = \pm 1$ dargestellt. Nach (11) ist nur noch $|x| \leq 2t - 2$ übrig. Dann ist $x^2 - py^2 \leq (2t-2)^2 - p(2u)^2 = -8t$ im Widerspruch mit (10). Somit ist die zweite Ungleichung in (11') zulässig. Gilt dann die erste nicht, so folgt $x^2 - py^2 \geq (2t)^2 - p(2u-1)^2 = 4pu - p - 4$, also nach (10) $4pu - p - 4 < 8t$ und nach (13) $4pu - p - 4 < 8u\sqrt{p}$. Wegen $u \geq \frac{3}{2}$ ergibt das $5p - 4 < 12\sqrt{p}$. Dieser Widerspruch beweist die Behauptung.

7. Das in **f** in Verbindung mit **a**^{IV} enthaltene Resultat sprechen wir in folgendem Lemma aus:

g. Existiert der E.A. im Fall I mit $p \geq 61$, so läßt sich zu jeder ganzen Körperzahl α eine zweite ganze Körperzahl ω so angeben, daß für die erste und zweite Koordinate²⁰⁾ und Norm von $\alpha + (\varepsilon + 1)\omega$

$$(15) \quad |(\alpha + (\varepsilon + 1)\omega)_1| < t.$$

$$(16) \quad |(\alpha + (\varepsilon + 1)\omega)_2| < u,$$

$$(17) \quad |N(\alpha + (\varepsilon + 1)\omega)| < 2t$$

²⁰⁾ Die Koordinaten einer Zahl $\gamma = x + y\sqrt{p}$ (x, y rational) werden durch $(\gamma)_1 = x$, $(\gamma)_2 = y$ definiert und bezeichnet. Die rechten Seiten in (15) bis (17) sind übrigenfalls $(\varepsilon)_1$, $(\varepsilon)_2$ und die Spur von ε .

gelten. Da es bloß auf die Restklasse $\alpha \pmod{\varepsilon + 1}$ ankommt, sind alle Möglichkeiten schon erschöpft, wenn man für nicht ganzes t $\alpha = r$ ($r = 0, 1, \dots, 2t - 1$), für ganzes t aber $\alpha = r$ und $\alpha = r + \frac{1 + \sqrt{p}}{2}$ ($r = 0, 1, \dots, t - 1$) setzt.

Aus dem auf $\mu = \varepsilon + 1$ angewendeten a^{IV} folgt nämlich, daß es ein $\bar{\mu}$ mit $\bar{\mu} \equiv \alpha \pmod{\varepsilon + 1}$ gibt. Wegen der Annahme ist jetzt nach¹⁹⁾ $u \geq \frac{3}{2}$, und somit kommt f' zur Geltung. Setzt man gleichzeitig $\bar{\mu} = \alpha + (\varepsilon + 1)\omega$, so entsteht g .

Der späteren Anwendung halber bemerken wir im Zusammenhang mit g folgendes:

g' . Ist α rational, so lautet (17)

$$(18) \quad |\alpha^2 + 2\alpha((\varepsilon + 1)\omega)_1 + 2tN\omega| < 2t,$$

und wenn auch (15) gilt, so ist

$$(19) \quad |N\omega| < |\alpha| + 1 + \frac{\alpha^2}{2t},$$

beides auch dann, wenn α und ω nicht ganz sind.

Die Richtigkeit folgt aus (12).

8. Zur Anwendung von g machen wir noch eine Vorbereitung:

h. Im Fall I mit $p > 5$ kann es in einem System S assoziierter ganzer Körperzahlen und ihrer Konjugierten bis auf das Vorzeichen höchstens nur einen Repräsentanten β geben, wofür wenigstens eine der Ungleichungen

$$(20) \quad |((\varepsilon + 1)\beta)_2| < \frac{t-1}{2},$$

$$(21) \quad |((\varepsilon + 1)\varepsilon'\beta')_2| < \frac{t-1}{2}$$

gilt. Ist S insbesondere die Gesamtheit der Einheiten, so gelten (20) und (21) nur mit $|\beta| = 1$, und zwar ist

$$(22) \quad |(\varepsilon + 1)_2| = |((\varepsilon + 1)\varepsilon')_2| = u.$$

Im allgemeinen Fall läßt sich die Beantwortung der Frage, ob (20) oder (21) oder beide gelten, durch den Zusammenhang

$$(23) \quad ((\varepsilon + 1)\beta)_2 + ((\varepsilon + 1)\varepsilon'\beta')_2 = 2(\beta)_2$$

erleichtern.

Wenn S aus den Einheiten oder aus den Assoziierten einer (ganz) rationalen Zahl besteht, ist alles klar, weshalb wir davon im folgenden absehen.

Wir setzen mit rationalen a_n, b_n

$$(24) \quad (\varepsilon + 1)\varepsilon'^n = a_n - b_n(\varepsilon + 1) \quad (n = 0, 1, 2, \dots)$$

und zeigen zuerst, daß a_n, b_n ganz sind und daß weiter

$$(25) \quad a_1 = 2t, \quad b_1 = 1,$$

$$(26) \quad \left(2t + 1 + \frac{1}{t}\right)b_n - 1 + \frac{1}{t} \geq a_n > (2t + 1)b_n > 0 \quad (n = 2, 3, \dots)$$

gelten. Man gewinnt nämlich durch Multiplizieren von (24) mit $\varepsilon' = 2t - \varepsilon$:

$$a_{n+1} = (2t + 1)a_n - 2tb_n, \quad b_{n+1} = a_n - b_n.$$

Wegen $a_0 = 0, b_0 = -1$ folgt hieraus (25) und $a_2 = 4t^2, b_2 = 2t - 1$. Also gilt (26) zuerst für $n = 2$, dann aber durch Induktion auch allgemein nach der Rechnung ($n \geq 2$):

$$\begin{aligned} a_{n+1} - (2t + 1)b_{n+1} &= b_n, \\ \left(2t + 1 + \frac{1}{t}\right)b_{n+1} - a_{n+1} &= \frac{1}{t}a_n - \left(1 + \frac{1}{t}\right)b_n > b_n > 1 - \frac{1}{t} \end{aligned}$$

Da nunmehr mit β auch $\varepsilon'\beta'$ zu S gehört, so dürfen wir annehmen, daß eben (20) gilt. Dann leiten wir aus

$$(27) \quad |((\varepsilon + 1)\varepsilon^n\beta)_2| < \frac{t-1}{2} \quad (n \neq 0; n \text{ ganz rational})$$

einen Widerspruch ab. Daraus wird natürlich auch folgen, daß (21) und

$$(28) \quad |((\varepsilon + 1)\varepsilon^n\beta')_2| < \frac{t-1}{2} \quad (n \neq 1; n \text{ ganz rational})$$

nicht gleichzeitig gelten können. Wenn wir also dann noch zeigen, daß auch (28) zu (20) im Widerspruch steht, so werden wir den vollen Beweis erbracht haben.

Wir dürfen in (27) $n > 0$ annehmen, da man sonst mit $\beta^* = \varepsilon'^n\beta$ statt β schließen könnte, wobei (20) und (27) vertauscht werden. Nach (24) und (27) folgt

$$|(a_n\beta - b_n(\varepsilon + 1)\beta)_2| < \frac{t-1}{2},$$

also wegen $b_n > 0$:

$$a_n |(\beta)_2| < b_n |((\varepsilon + 1)\beta)_2| + \frac{t-1}{2}.$$

Da β nicht rational ist, so ist $|(\beta)_2| \geq \frac{1}{2}$, und somit ergibt sich nach (20), (25), (26)

$$\frac{1}{2}a_n < \frac{t-1}{2}(b_n + 1), \quad tb_n < \frac{t-1}{2}(b_n + 1).$$

Das ist aber ein Widerspruch, da $b_n \geq 1$ ist.

Es ist noch (28) zu widerlegen. Infolge (20) müssen die Koordinaten von β (von Null verschieden und) von entgegengesetztem Vorzeichen sein. Dann ist in (28) sicher $n > 0$, also $n \geq 2$. Nach Einsetzen von (24) und Übergang zum Konjugierten ergibt sich

$$|((a_n - b_n(\varepsilon' + 1))\beta)_2| < \frac{t-1}{2}.$$

Der mit β multiplizierte Teil läßt sich wieder wegen $\varepsilon' = 2t - s$ in der Form

$$(-a_n + (2t + 2)b_n)\varepsilon + (a_n - (2t + 1)b_n)(\varepsilon + 1)$$

schreiben. Nach (26) folgt hieraus

$$(-a_n + (2t + 2)b_n)|(\varepsilon\beta)_2| < (a_n - (2t + 1)b_n)|((\varepsilon + 1)\beta)_2| + \frac{t-1}{2}.$$

Es ist wieder $|(\varepsilon\beta)_2| \geq \frac{1}{2}$, weiter ist nach (20) offenbar $|((\varepsilon + 1)\beta)_2| \leq \frac{t}{2} - \frac{3}{4}$, also

$$\frac{1}{2}(-a_n + (2t + 2)b_n) < (a_n - (2t + 1)b_n)\left(\frac{t}{2} - \frac{3}{4}\right) + \frac{t-1}{2},$$

$$(4t^2 + 1)b_n < (2t - 1)a_n + 2t - 2.$$

Nach (26) folgt

$$(4t^2 + 1)b_n < \left(4t^2 - 1 + \frac{2t-1}{t}\right)b_n + \frac{t-1}{t}$$

und weiter hieraus der Widerspruch $b_n < t - 1$. Damit haben wir **h** bewiesen.

9. Wir beweisen, daß der E. A. in $K = R(\sqrt{61})$ nicht existiert. Betrachten wir vorläufig ein beliebiges $K = R(\sqrt{p})$ mit $p \equiv 5 \pmod{8}$, $p \geq 61$ und nicht ganzem t . Dann sind $2t$, $2u$ ungerade ganz, letzteres $\equiv 1 \pmod{4}$. Nehmen wir an, daß der E. A. existiert. Es werde in **g** $\alpha = t + \frac{1}{2}$ gesetzt, das jetzt ja ganz ist. Eine beliebige ganze Zahl ω in K läßt sich durch

$$(29) \quad \omega = \frac{1}{2}(-(\varepsilon' + 1) + \beta),$$

$$(30) \quad \beta = \frac{r+s\sqrt{p}}{2} \quad (r, s \text{ ganz rational}),$$

$$(31) \quad r \equiv s \equiv 1 \pmod{2}, \quad r + s \equiv 2t + 1 \pmod{4}$$

angeben. Wegen (12) und (29) ist

$$(32) \quad \alpha + (\varepsilon + 1)\omega = \frac{1}{2} + \frac{1}{2}(\varepsilon + 1)\beta.$$

Nach **g** läßt sich also ein solches (ω d. h.) β wählen, daß [nach Beachtung von **g'**, wobei jetzt gemäß (32) $\alpha = \frac{1}{2}$, $\omega = \frac{\beta}{2}$ zu setzen ist]:

$$(33) \quad \frac{1}{2}|((\varepsilon + 1)\beta)_2| < u,$$

$$(34) \quad \left| \frac{1}{4} + \frac{1}{2}((\varepsilon + 1)\beta)_1 + \frac{t}{2}N\beta \right| < 2t,$$

$$\frac{1}{4}|N\beta| < \frac{1}{2} + 1 + \frac{1}{8t}.$$

Nach dem letzteren ist $|N\beta| \leq 6$. Wäre β eine Einheit, so müßte nach (33) und **h** $\beta = \pm 1$ oder $\pm \varepsilon'$ sein; beide führen zu keinem ganzen ω in (29). Da weiter nach (30) und (31) $N\beta$ ungerade ist, ist $|N\beta| = 3$ oder 5.

Nunmehr sei $p = 61$, $\varepsilon = \frac{39+5\sqrt{61}}{2}$. Dann wird (33) durch die ganzen Zahlen

$$\beta_3 = \pm (8 - \sqrt{61}), \quad \beta_5 = \pm \frac{9+\sqrt{61}}{2}$$

(es ist $N\beta_i = i$) befriedigt. Nach h kann das gleiche wegen $2u < \frac{t-1}{2}$ nur noch über $\varepsilon' \beta'_i$ ($i = 3, 5$), d. h. über

$$\beta_3^* = \pm \frac{7-\sqrt{61}}{2}, \quad \beta_5^* = \pm \frac{23-3\sqrt{61}}{2}$$

gelten. Von diesen acht Zahlen genügen (31) nur β_5, β_5^* . Für diese gilt aber (34) nicht, da das dritte Glied $\pm \frac{5t}{2}$, das mittlere ± 8 bzw. $\pm \frac{7}{2}$ ist. Damit haben wir die Behauptung über $p = 61$ bewiesen²¹⁾.

10. Um die Nichtexistenz des E. A. in $K = R(\sqrt{109})$ zu beweisen, nehmen wir jetzt für α in g die ganz rationale Zahl mit $\sqrt{2t} \leq \alpha < \sqrt{2t} + 1$. Dann ist $\omega \neq 0$, und (16) geht in

$$(35) \quad |((\varepsilon + 1)\omega)_2| < u$$

über. Weiter ist nach (19), da ω ganz ist: $|N\omega| \leq \alpha + 2$.

Ist nun $p = 109$, $\varepsilon = \frac{261+25\sqrt{109}}{2}$, so ist $\alpha = 17$, $|N\omega| \leq 19$. Wegen (35) und h kann ω keine Einheit, noch das Assoziierte einer rationalen Zahl sein. Es gibt bis auf Assoziierte und Konjugierte nur drei nichtrationale Primzahlen $\omega_3, \omega_5, \omega_7$ (es ist $N\omega_i = i$), deren Norm absolut nicht größer als 19 ist. Ein zusammengesetztes ω kann also bis auf Assoziierte und Konjugierte nur $2\omega_3, \omega_3^2, \omega_3\omega_5$ oder $\omega'_3\omega_5$ sein. Wählt man

$$\omega_3 = \pm \frac{11-\sqrt{109}}{2}, \quad \omega_5 = \pm (21 - 2\sqrt{109}), \quad \omega_7 = \pm (94 - 9\sqrt{109}),$$

so wird (35) durch diese Primzahlen und auch durch

$$\omega_{12} = 2\omega_3, \pm \omega_9 = \pm \omega_3^2 = \pm \frac{115-11\sqrt{109}}{2},$$

$$\omega_{15} = \omega_3\omega_5 = \pm \frac{449-43\sqrt{109}}{2}, \quad \omega_{15}^* = \omega'_3\omega_5 = \pm \frac{13-\sqrt{109}}{2}$$

²¹⁾ Es läßt sich zeigen, daß die H_Q für $p = 61$ alle Punkte des Fundamentalbereiches T bis auf das Rechteck mit den Seiten $x = 0$, $x = 0,0203$, $y = 0,216199$, $y = 0,219036$ überdecken. Nach den Überlegungen von Remak [15], S. 247 erhält man daraus durch einfache Rechnung, daß unter den rationalen Punkten von T höchstens nur $(\frac{1}{78}, \frac{17}{78})$ unüberdeckbar sein kann. Daß dieser Punkt wirklich unüberdeckbar ist,

folgt aus dem Obigen. Auch müssen $\frac{1+17\sqrt{61}}{78}$ und $(\frac{\alpha}{\varepsilon+1}) = \frac{2t+1}{2(\varepsilon+1)}$ äquivalent sein; sie sind sogar kongruent mod 1. Das spricht dafür, daß obiger Beweis im wesentlichen der einzig mögliche ist. Vgl.¹¹⁾.

befriedigt (die Stellenzeiger geben wieder die Norm an). Dann kann ω nach h außer diesen 14 Zahlen nur noch

$$\varepsilon' \omega_3 = \pm \frac{23-7\sqrt{109}}{2}, \quad \varepsilon' \omega_5 = \pm \frac{31-3\sqrt{109}}{2}, \quad \varepsilon' \omega_7 = \pm \frac{9-\sqrt{109}}{2},$$

$$2\varepsilon' \omega_3, \quad \varepsilon' \omega_9 = \pm (10 - \sqrt{109}), \quad \varepsilon' \omega_{15} = \pm \frac{7-\sqrt{109}}{2},$$

$$\varepsilon' \omega_{15}^* = \pm (167 - 16\sqrt{109})$$

sein. Man rechnet aber leicht nach, daß keine dieser 28 Zahlen (18) befriedigt, das jetzt nämlich

$$|289 + 17((263 + 25\sqrt{109})\omega)_1 + 261N\omega| < 261$$

lautet. Mithin gilt der E.A. für $p = 109$ nicht.

11. Wir betrachten jetzt den Fall I mit $p \equiv 17 \pmod{24}$, $p > 41$ und beweisen, daß dann der E.A. nicht existiert. Jetzt ist

$$(36) \quad t \equiv 0 \pmod{4}, \quad u \equiv 1 \pmod{2}.$$

Nehmen wir an, daß der E.A. existiert. Da $p \geq 89$, können wir g anwenden. Das tun wir mit $\alpha = \frac{s+1}{2} + 3$; wegen (36) ist ja α ganz. Gleichzeitig setzen wir in $g \omega = -\frac{1}{2} + \frac{\beta}{2}$, wobei β eine ganze Zahl in K mit

$$(37) \quad \beta \equiv 1 \pmod{2}$$

ist; (37) ist notwendig und hinreichend, damit ω ganz ist. Wegen $\alpha + (s+1)\omega = 3 + \frac{1}{2}(s+1)\beta$ folgt, indem man auch g' beachtet (nämlich mit $\alpha = 3$, $\omega = \frac{\beta}{2}$), daß es ein β gibt mit

$$(38) \quad |6 + ((s+1)\beta)_1| < 2t,$$

$$(39) \quad |((s+1)\beta)_2| < 2u,$$

$$(40) \quad |18 + 6((s+1)\beta)_1 + tN\beta| < 4t,$$

$$(41) \quad |N\beta| < 16 + \frac{18}{t}.$$

Nach der Einleitung kann nur $p = 89, 113, 137$ oder $p > 3001$ sein. Da für die ersten drei Fälle

$$(42) \quad s = 500 + 53\sqrt{89}, \quad 776 + 73\sqrt{113} \quad \text{bzw.} \quad 1744 + 149\sqrt{137}$$

ist, so ist immer $t > \sqrt{3000}$, also

$$(43) \quad t > 54.$$

[Nach (36) könnte man etwas mehr sagen.]

Hiernach folgt aus (41) $|N\beta| \leq 16$, sogar $|N\beta| \leq 15$, da $|N\beta|$ ungerade ist. Es kann β keine Einheit sein. Wegen (39) und h (es ist offenbar $\frac{t-1}{2} \geq 2u$) kommen nämlich bloß $\beta = \pm 1$, $\beta = \pm \varepsilon'$ in Frage. Dann ist $((\varepsilon+1)\beta)_1 = \pm(t+1)$ bzw. $\pm(t-1)$. Beide sind aber mit (40), (43) und $N\beta = \pm 1$ unvereinbar. Wegen (39) und h kann β auch nicht assoziiert zu einer rationalen Zahl sein. Auch muß $(\beta, 3) = 1$ sein. Da nämlich 3 eine Primzahl in K ist, würde aus $(\beta, 3) \neq 1$ folgen, daß $3|\beta$, $9|N\beta$, $|N\beta| = 9$ ist, und dann wäre β doch zu 3 assoziiert. Nach all diesem kann nur

$$(44) \quad |N\beta| = 5, 7, 11, 13$$

sein.

Dann sind die Koordinaten von β von Null verschieden und wegen (39) von entgegengesetztem Vorzeichen. Sie sind wegen (37) auch ganz. Wir können also setzen

$$(45) \quad \beta = \sigma(x - y\sqrt{p}) \quad (x, y \text{ positiv ganz, } 2 \nmid x+y, \sigma = \pm 1).$$

Setzen wir noch der Kürze halber

$$(46) \quad b = N\beta.$$

Dann ist

$$(47) \quad py^2 = x^2 - b.$$

Betrachten wir zuerst $p > 3001$ (wir lassen also $p = 89, 113, 137$ vorläufig beiseite). Dann gilt

$$(48) \quad x > 54.$$

Wir setzen

$$(49) \quad x^* = tx - puy.$$

Da $\pm x^*$ die erste Koordinate von $\varepsilon\beta$ ist, so folgt ebenfalls

$$(50) \quad |x^*| > 54.$$

Wegen (6) und (47), (48) ist

$$(51) \quad t < u\sqrt{p} < t + \frac{1}{2t},$$

$$(52) \quad x - \frac{7}{x} < y\sqrt{p} < x + \frac{7}{x}.$$

Nach (45) ist

$$(53) \quad ((\varepsilon+1)\beta)_1 = \sigma((t+1)x - puy).$$

Wir zeigen, daß

$$(54) \quad (t+1)x - puy > 0$$

ist. Nach (53), (40) und $|b| \geq 5$ ist nämlich

$$|(t+1)x - puy| \geq \frac{t-13}{6}.$$

Wäre (54) falsch, so wäre

$$(t+1)x + \frac{t-18}{6} \leq puy.$$

Nach Quadrieren folgt, da nach (47) $py^2 \leq x^2 + 13$ ist,

$$(t+1)^2 x^2 + (t+1)(t-18) \frac{x^2}{3} < p^2 u^2 y^2 \leq (t^2+1)(x^2+13),$$

$$6tx^2 + (t+1)(t-18)x < 39(t^2+1),$$

$$6x^2 + (t-18)x < 39t + \frac{39}{t}.$$

Andererseits ist nach (43) und (48) $6x^2 - 18x > \frac{39}{t}$, also $tx < 39t$, $x < 39$.

Dieser Widerspruch mit (48) beweist (54).

Nach (53), (54), (38), (40) ist

$$(55) \quad \frac{(|b|-4)t-18}{6} < (t+1)x - puy < 2t+6.$$

Nunmehr wollen wir der Reihe nach die Unmöglichkeit von $b=5$, $b=-5$, $|b|>5$ ausweisen.

Fall $b=5$. Aus (47), (48) folgt

$$(56) \quad x - \frac{3}{x} < y\sqrt{p} < x.$$

Hieraus und aus (51) gewinnen wir

$$(57) \quad (t+1)x - puy > (t+1)x - \left(t + \frac{1}{2t}\right)x = \frac{2t-1}{2t}x.$$

Nach (55) folgt weiter $x < \frac{2t}{2t-1}(2t+6)$, also nach (43)

$$(58) \quad x \leq 2t+7.$$

Ebenfalls aus (51) und (56) folgt

$$(59) \quad (t+1)x - puy < (t+1)x - t\left(x - \frac{3}{x}\right) = x + \frac{3t}{x}$$

und weiter aus (55)

$$x + \frac{3t}{x} > \frac{t-18}{6}.$$

Wäre hier $\frac{3t}{x} > x$, so würde $\frac{6t}{x} > \frac{t-18}{6}$, $x < \frac{36t}{t-18}$, also nach (43) $x < \frac{36 \cdot 54}{36} = 54$ folgen, im Widerspruch mit (48). Somit ist $x \geq \frac{3t}{x}$,

$$(60) \quad x > \frac{t-18}{12}.$$

Nach (49) und (57) ist $x^* > -\frac{x}{2t}$. Hieraus folgt nach (58)

$$(61) \quad x^* > -2.$$

Nach (49) und (59) ist $x^* < \frac{3t}{x}$. Hieraus folgt nach (60) $x^* < \frac{36t}{t-18}$, also nach obiger Auswertung $x^* < 54$. Dies und (61) ergeben einen Widerspruch mit (50).

Fall $b = -5$. Aus (47) folgt

$$(62) \quad x < y\sqrt{p} < x + \frac{5}{2x}.$$

Hieraus und aus (51) gewinnen wir

$$(63) \quad (t+1)x - puy > (t+1)x - \left(t + \frac{1}{2t}\right)\left(x + \frac{5}{2x}\right) > \frac{2t-1}{2t}x - (t+1)\frac{5}{2x},$$

also aus (43)

$$(t+1)x - puy > \frac{2t-1}{2t}x - \frac{t+1}{20}.$$

Dies, (55) und (43) ergeben

$$(64) \quad x < \frac{2t}{2t-1}(2t+6) + \frac{2t}{2t-1} \frac{t+1}{20} < \frac{41t}{20} + 8 = \frac{41t+160}{20}.$$

Wieder aus (51) und (62) folgt

$$(65) \quad (t+1)x - puy < (t+1)x - tx = x$$

und weiter aus (55)

$$(66) \quad x > \frac{t-18}{6}.$$

Nach (49) und (63) ist $x^* > -\frac{x}{2t} - (t+1)\frac{5}{2x}$, also nach (64) und (66)

$$x^* > -\frac{41t+160}{40t} - \frac{15(t+1)}{t-18}.$$

Hieraus folgt nach (43)

$$(67) \quad x^* > -25.$$

Nach (49) und (65) ist $x^* < 0$, und das ergibt mit (67) und (50) einen Widerspruch.

Fall $|b| > 5$. Dann ist $7 \leq |b| \leq 13$. Nach (55) ist jetzt

$$(68) \quad \frac{t-6}{2} < (t+1)x - puy < 2t+6.$$

Es ist nach (51) und (52)

$$(69) \quad (t+1)x - puy > (t+1)x - \left(t + \frac{1}{2t}\right)\left(x + \frac{7}{x}\right) > \frac{2t-1}{2t}x - (t+1)\frac{7}{x},$$

also nach (48)

$$(t+1)x - pux > \frac{2t-1}{2t}x - \frac{t+1}{7}.$$

Hieraus und aus (43), (68) folgt

$$(70) \quad x < \frac{2t}{2t-1}(2t+6) + \frac{2t}{2t-1} \frac{t+1}{7} < \frac{15t}{7} + 9.$$

Nach (51) und (52) ist

$$(71) \quad (t+1)x - puy < (t+1)x - t \left(x - \frac{7}{y}\right) = x + \frac{7t}{x},$$

also nach (68)

$$x + \frac{7t}{x} > \frac{t-6}{2}.$$

Wieder kann nicht $\frac{7t}{x} > x$ sein, denn dann wäre $\frac{14t}{x} > \frac{t-6}{2}$, im Widerspruch mit (43), (48). Also ist $x \geq \frac{7t}{x}$,

$$(72) \quad x > \frac{t-6}{4}.$$

Nach (49) und (69) ist

$$x^* > -\frac{x}{2t} - (t+1)\frac{7}{x},$$

woraus weiter nach (70), (72) und (43) $x^* > -\frac{15t+63}{14t} - \frac{28(t+1)}{t-6}$,

$$(73) \quad x^* > -35.$$

Wegen (49), (71), (72) und (43) ist $x^* < \frac{7t}{x} < \frac{28t}{t-6} < 32$. Dies und (73) widersprechen (50), womit der Beweis für $p > 3001$ erbracht ist.

Es sind noch $p = 89, 113, 137$ übrig (man beachte (42)!).

Fall $p = 89$. Wegen (44) ist nur $|N\beta| = 5, 11$ möglich. Ähnlich wie in 9. und 10. sehen wir, daß wegen (39) und (42) für β nur

$$\beta_5 = \pm (19 - 2\sqrt{89}), \quad \varepsilon'\beta'_5 = \pm (66 - 7\sqrt{89}), \quad \beta_{11} = \pm (10 - \sqrt{89}),$$

$$\varepsilon'\beta'_{11} = \pm (283 - 30\sqrt{89})$$

möglich sind. Keines von ihnen genügt (40).

Fall $p = 113$. Es ist $|N\beta| = 7, 11$ oder 13. Jetzt kämen

$$\beta_7 = \pm (32 - 3\sqrt{113}), \quad \varepsilon'\beta'_7 = \pm (85 - 8\sqrt{113}), \quad \beta_{11} = \pm (202 - 19\sqrt{113}),$$

$$\varepsilon'\beta'_{11} = \pm (21 - 2\sqrt{113}), \quad \beta_{13} = \pm (489 - 46\sqrt{113}), \quad \varepsilon'\beta'_{13} = \pm (10 - \sqrt{113})$$

in Betracht, ohne jedoch (40) zu genügen.

Fall $p = 137$. Dann ist $|N\beta| = 7$ oder 11. Es ist ähnlich zu schließen mit

$$\beta_7 = \pm (12 - \sqrt{137}), \quad \varepsilon'\beta'_7 = \pm (515 - 44\sqrt{137}), \quad \beta_{11} = \pm (82 - 7\sqrt{137}),$$

$$\varepsilon'\beta'_{11} = \pm (117 - 10\sqrt{137}).$$

Damit ist die Behauptung bewiesen.

* * *

Nachwort. Man darf nicht glauben, daß man jetzt, da man im Fall I die Restklassen $p \equiv 5, 13, 17 \pmod{24}$ schon hinter sich und nur noch $p \equiv 1 \pmod{24}$ übrig hat, das Problem zu „drei Viertel“ gelöst hätte, da letztere p viel mehr Schwierigkeiten zu bereiten scheinen, gleichgültig, ob man durch h oder g ans Werk gehen will; und sie steigern sich, je größer die kleinste Primzahl a (> 0) mit $\left(\frac{a}{p}\right) = -1$ ist. Ähnliches zeigte sich auch im Fall II, erst bei Rédei [13] (und [14]) haben sich alle Unterschiede zwischen den verschiedenen Restklassen so gut wie verwischt.

Die Primzahlen mit E. A. in den obigen drei Restklassen bilden die lückenlose Folge 5, 13, 17, 29, 37, 41, worin alle drei mit je zwei Elementen vertreten sind. Erst nach der Lücke 53, 61 tritt dann endlich in der Restklasse 1 ($\pmod{24}$) die erste Primzahl 73 auf, ebenfalls mit E. A. Man vermutet mit Recht, daß das auch seine Fortsetzung hat, da schon unter 3001 die obigen 97, 193, ..., 601 noch in Frage kommen (das Nachprüfen für 97 scheint nicht sehr schwer zu sein).

Zu $p \equiv 17 \pmod{24}$, $109 < p < 3001$ haben wir h gebraucht. Diese Hilfe läßt sich wahrscheinlich ausschalten (in 11. haben wir zu grob abgeschätzt!). Es wäre erwünscht, Brauers [3] Resultat für $p \equiv 13 \pmod{24}$, $p > 109$ auch durch g abzuleiten, wobei man dann erwartet, daß auch $109 < p < 3300000$ (wenigstens zum größten Teil) keiner Sonderbehandlung bedarf.

Bemerkung nach der Korrektur. Auch $R(\sqrt{97})$ ist mit E. A., wie sich inzwischen herausstellte.

Literaturverzeichnis.

- [1] H. Behrbohm u. L. Rédei, Der Euklidische Algorithmus in quadratischen Zahlkörpern. Journ. f. d. Math. 174 (1936), S. 192–205.
- [2] E. Berg, Über die Existenz eines Euklidischen Algorithmus in quadratischen Zahlkörpern. Kungl. Fysiografiska Sällskapets i Lund Förhandlingar 5 (1935), Nr. 5.
- [3] A. Brauer, On the non-existence of the Euclidean algorithm in certain quadratic number fields. Amer. J. Math. 62 (1940), S. 697–716.
- [4] L. E. Dickson, Algebren und ihre Zahlentheorie. Zürich u. Leipzig (1927), S. 150–151.
- [5] L. Dirichlet u. R. Dedekind, Vorlesungen über Zahlentheorie, 4. Aufl., Braunschweig (1894), S. 451.
- [6] P. Erdős u. Ch. Ko, Note on the Euclidean algorithm. Journ. of the London Math. Soc. 13 (1938), S. 3–8.
- [7] J. Fox Keston, Existence of a Euclidean algorithm in quadratic fields. Thesis Yale University (1935).
- [8] H. Heilbronn, On Euclid's algorithm in real quadratic fields. Proc. of the Cambridge Phil. Soc. 34 (1938), S. 521–526.
- [9] N. Hofreiter, Quadratische Körper mit und ohne Euklidischen Algorithmus. Monatsh. f. Math. u. Phys. 15 (1935), S. 397–460.

- [10] N. Hofreiter, Quadratische Zahlkörper ohne Euklidischen Algorithmus. *Math. Annalen* 110 (1935), S. 195–196.
- [11] A. Oppenheim, Quadratic fields with and without Euclid's algorithm. *Math. Annalen* 109 (1934), S. 349–352.
- [12] O. Perron, Quadratische Zahlkörper mit Euklidischem Algorithmus. *Math. Annalen* 107 (1932), S. 489–495.
- [13] L. Rédei, Über den Euklidischen Algorithmus in reellquadratischen Zahlkörpern. *Mat. Fiz. Lapok* 47 (1940), S. 78–90 (ungarisch).
- [14] L. Rédei, Über den Euklidischen Algorithmus in reellquadratischen Zahlkörpern. *Journ. f. d. Math.* 183 (1941), S. 183–192.
- [15] R. Remak, Über den Euklidischen Algorithmus in reell-quadratischen Zahlkörpern. *Jahresber. d. D. M. V.* 44 (1934), S. 238–250.
- [16] L. Schuster, Reellquadratische Zahlkörper ohne Euklidischen Algorithmus. *Monatsh. f. Math. u. Phys.* 47 (1938), S. 117–127.

(Eingegangen am 19. 8. 1941.)

Preisauflage der Technischen Hochschule Karlsruhe.

Die Technische Hochschule Fridericiana in Karlsruhe erläßt zum Gedächtnis des durch lange Jahre in ihrem Lehrkörper wirkenden Bahnbrechers der mathematischen Logik Ernst Schröder aus Anlaß der 100. Wiederkehr seines Geburtstages die folgende *Preisauflage*:

Die Untersuchungen Ernst Schröders zum Eliminations- und Auflösungsproblem in der Logik der Relative [Vorlesungen über die Algebra der Logik III, 1, Leipzig 1895; vgl. J. C. C. Mc Kinsey, Postulates for the calculus of binary relations, *J. Symbolic Logic* 3 (1940), S. 85–97] sollen mit den heute verfügbaren Hilfsmitteln [vgl. H. Behmann, Beiträge zur Algebra der Logik, insbesondere zum Entscheidungsproblem, *Math. Annalen* 86 (1922), S. 163–229, und vor allem W. Ackermann, Untersuchungen über das Eliminationsproblem der mathematischen Logik, ebenda 110 (1934), S. 390–314 und Zum Eliminationsproblem der mathematischen Logik, ebenda 111 (1935), S. 61–63, sowie Hilbert-Ackermann, Grundzüge der theoretischen Logik, 2. Aufl., Berlin 1938, S. 106f.] weitergeführt werden.

Als Preis wird der Betrag von 500 RM für die beste Lösung der Aufgabe ausgesetzt. Als Zeitpunkt für die Einlieferung der Lösungen wird mit Rücksicht auf die Kriegsverhältnisse vorläufig der 1. Oktober 1944 angesetzt. Dem Preisrichterkollegium gehören außer dem Rektor der Technischen Hochschule Karlsruhe als Vorsitzendem, Prof. Dr.-Ing. R. G. Weigel, und dem Dekan der math.-naturw. Fakultät Prof. Dr. A. Bühl an: Studienrat Dr. Wilh. Ackermann (Burgsteinfurt), Prof. Dr. C. Boehm (Karlsruhe), der Herausgeber der Frege-Studien Jürgen v. Kempski (Berlin), Prof. Dr. Heinr. Scholz (Münster i. W.) sowie Prof. Dr. E. Ungerer in Karlsruhe als Geschäftsführer. Etwaige Anfragen sind an den Geschäftsführer zu richten.

David Hilbert

† 14. Februar 1943

Am 14. Februar 1943 ist David Hilbert im Alter von 81 Jahren in Göttingen gestorben. Damit ist der letzte aus der zweiten großen Göttinger mathematischen Blütezeit dahingegangen. Bei einer solchen Nachricht geht auch in unserer ereignisreichen Kriegszeit ein ehrfürchtiges Gedenken an diesen mathematischen Genius durch die gesamte mathematische Welt, die länger als ein Menschenalter unter dem Bann der Offenbarungen seines Geistes gestanden hat.

Die Mathematischen Annalen sind mit Hilbert während seines ganzen Lebens auf das engste verbunden gewesen. Von seiner Doktorarbeit an [Math. Ann. Bd. 30 (1887)] bis zu seiner letzten Arbeit in Math. Ann. Bd. 104 (1931) hat Hilbert einen großen Teil seiner Arbeiten in den Annalen veröffentlicht, die er dann als Herausgeber seit Bd. 55 (1902) bis Bd. 116 (1939) geleitet und entscheidend in ihrer Haltung beeinflußt hat. Der Name Hilberts wird auch mit unserer Zeitschrift immer verknüpft sein.

Die Redaktion
der Mathematischen Annalen

Die Dualität von Inzidenz und Senkrechtstehen in der absoluten Geometrie.

Von

Arnold Schmidt in Marburg a. d. Lahn.

§ 1.

Einleitung und Übersicht.

Zur Begründung der ebenen absoluten Geometrie bedient man sich seit Hjelmslev¹⁾ in grundlegender Weise der Spiegelungen. Neben den Hjelmslevschen Arbeiten sind hier vor allem die Begründung der vollelliptischen Geometrie durch Podehl und Reidemeister²⁾ und diejenige der absoluten Geometrie durch Bachmann und Reidemeister³⁾ zu nennen, bei denen die axiomatische Grundlage noch etwas weiter (z. B. durch Vermeidung der Streckenhalbierungsforderung) eingeschränkt ist. In dem betreffenden Axiomensystem wird zwar die Streckenkongruenz als ein Grundbegriff zugrunde gelegt; ihre axiomatische Determination geht aber gerade nur so weit, wie zur Definition der Spiegelungen nötig ist. Es wird daher fruchtbar sein, statt der eingeschränkten Streckenkongruenz die Spiegelung direkt als Grundbegriff zu nehmen; dabei fallen außer der Kongruenz auch Inzidenz und Senkrechtstehen als Grundrelationen fort. (Im folgenden Axiomensystem sind der Einfachheit halber außer den Geradenspiegelungen auch Punktspiegelungen herangezogen, dies ist offenbar nicht nötig. Die Axiome werden übrigens die „absolute Geometrie“ gleich in jenem weiteren Sinne aufspannen, der die Frage nach der Existenz der Pole offenläßt und somit den vollelliptischen Fall und den im engeren Sinne absoluten Fall, in dem Pole verboten sind⁴⁾, gemeinsam umschließt.) — Ein Axiomensystem für die *euklidische* Geometrie (und ansatz-

¹⁾ Neue Begründung der ebenen Geometrie. Math. Annalen 64 (1907), ab S. 449 und Einleitung in die allgemeine Kongruenzlehre. Math. fys. Meddelelser, Kgl. Danske Videnskabskabernes Selskab 8 (1929), ab S. 11 und 10 (1929), ab S. 1.

²⁾ Eine Begründung der elliptischen Geometrie. Hamburger Abh. 10 (1934), ab S. 231.

³⁾ (1) F. Bachmann, Eine Begründung der absoluten Geometrie in der Ebene. Math. Annalen 113 (1936), ab S. 424. (2) F. Bachmann und K. Reidemeister, Die metrische Form in der absoluten und der elliptischen Geometrie. Math. Annalen 113 (1937), ab S. 748.

⁴⁾ Die in Anm. 2 genannte Arbeit bezieht sich auf den ersteren Fall, die in Anm. 3 (1) genannte Arbeit auf den letzteren.

weise für die hyperbolische), das sich der Punkt- und Geradenspiegelungen als Grundbegriffe bedient, wurde bekanntlich von Thomsen⁵⁾ gegeben; der Blickrichtung einer gruppentheoretischen Allgemeinheit, die dort für die Formulierung der Axiome maßgebend war, steht hier das Bestreben gegenüber, der aus dem genannten einzigen Grundbegriff sich entwickelnden absoluten Geometrie — die im folgenden kurz als „reine Spiegelungsgeometrie“ bezeichnet werden möge — eine anschaulich möglichst übersichtliche und handliche Axiomenbasis zugrunde zu legen.

Hierbei ergibt sich nun — § 2 — eine bemerkenswerte *Dualität der Axiome, eine Dualität zwischen Inzidenz und Senkrechtstehen*. Die explizite Herausstellung dieser Dualität verleiht nicht nur dem Axiomensystem eine augenfällige Geschlossenheit, sie ist vielmehr auch geeignet, die Unterteilung der absoluten Geometrie in ihre wichtigen Unterfälle zu systematisieren sowie den deduktiven Aufbau zu vereinfachen. Sie führt weiter im Falle der voll-elliptischen Geometrie zu einem aus nur ganz wenigen Axiomen bestehenden Axiomensystem (dem als einziges Grundgebilde das Aggregat „Gerade nebst Pol“ zugrunde liegt).

Bei der „Begründung“ (d. h. hier im wesentlichen: der projektiven Erweiterung und der auf dieser fußenden Metrisierung) der „reinen Spiegelungsgeometrie“ kann man verschiedene Gesichtspunkte zur Geltung bringen. Sie ließe sich z. B. ähnlich wie diejenige durchführen, die Bachmann und Reide-meister auf Grund ihres Axiomensystems gaben, bei der der Desarguessche Satz durch eine Heranziehung des kinematischen Raumes gewonnen wird und sodann das Vorhandensein einer linearen involutorischen Polarkorrelation an Hand einer zum Teil algebraischen Überlegung bewiesen wird. Da die Möglichkeit einer strengen und unmittelbaren projektiven Erweiterung von der erwähnten engen axiomatischen Basis aus ohne Durchgang durch einen Bewegungsraum bislang nicht aufgezeigt wurde, mag die in § 3 gegebene projektive Erweiterung, die das wirkliche Ausreichen des Axiomensystems zeigen soll, diesen Gesichtspunkt als durchführbar herausstellen. — Nach Erlangung eines Erweiterungsbereichs, der den projektiven Grundgesetzen genügt, wird auf die Gegenpaarung zurückgegriffen, welche sich auch z. B. bei der Metrisierung als besonders durchschlagend erweist. Bei dem Beweise — § 4 —, daß die Polarkorrelation, die im nichteuklidischen Falle die Bewegungsgruppe fixiert, linear ist (synthetisch ausgedrückt: daß sie auf zwei Geraden projektive Involutionen ausschneidet), gestattet nämlich die Gegenpaarung in einfacher Weise synthetisch, ohne Rechnung, zu schließen.

⁵⁾ Grundlagen der Elementargeometrie in gruppentheoretischer Behandlung. Hamburger math. Einzelschr. 15 (1933).

§ 2.

Die in sich duale Axiomenbasis.

Die Axiome beschäftigen sich mit zwei (nicht leeren⁶⁾ und nicht notwendig getrennten) Dingklassen — den Punkten A, B, \dots und den Geraden a, b, \dots — und mit einer einzigen Verknüpfung zwischen den Dingen und ihren Verknüpfungsreihen. Diese Verknüpfung soll stets ausführbar sowie assoziativ und gleichheitstreu sein und der „Streichungsregel“ genügen: „Es ist stets $XX\mathfrak{A}\mathfrak{B} = \mathfrak{A}\mathfrak{B}$ und ebenso $\mathfrak{A}xx\mathfrak{B} = \mathfrak{A}\mathfrak{B}$ “; hier darf eine der beiden durch $\mathfrak{A}, \mathfrak{B}$ mitgeteilten Verknüpfungsreihen die leere Reihe sein. (Auf Grund der Streichungsregel stimmt eine Dingverknüpfungsreihe, die einem Punkt oder einer Geraden gleich ist, mit ihrer Umkehrung überein.)

Die Relation „ $\mathfrak{A}\mathfrak{B} = \mathfrak{B}\mathfrak{A}$ bei $\mathfrak{A} \neq \mathfrak{B}$ “ sei durch $\mathbf{K}\mathfrak{A}\mathfrak{B}$ abgekürzt.

Interpretation. Eine Gerade vertritt die Spiegelung an ihr, ein Punkt die Umwendung an ihm, die Verknüpfung bedeutet die Zusammensetzung solcher Bewegungen. Man hat hiermit die Deutungen:

$\mathbf{K}pq$: p, q stehen senkrecht aufeinander. pq ist dann der Schnittpunkt.

$\mathbf{K}PQ$: P, Q sind konjugiert. PQ ist dann die Verbindung.

$\mathbf{K}pQ$ (oder auch $\mathbf{K}Qp$): p geht durch Q . pQ ist dann das Lot auf p in Q .

$p = Q$: p und Q stehen in Pol-Polaren-Beziehung. pQ ist dann die Identität (für die kein besonderes Zeichen eingeführt zu werden braucht, vgl. die Streichungsregel).

Die Deutung der folgenden Axiome ergibt sich hieraus unmittelbar.

Axiome der Spiegelungsgeometrie.

(Die dualisierten Elemente sind durch die Benennungen x, y, X, Y hervorgehoben.)

- | | |
|---------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|
| 1a. Zu P, X gibt es a mit $\mathbf{K}aP, \mathbf{K}aX$. | 1b. Zu P, x gibt es a mit $\mathbf{K}aP, \mathbf{K}ax$. |
| 2a. Bei $\mathbf{K}aX$ gibt es $y = aX$. | 2b. Bei $\mathbf{K}ax$ gibt es $Y = ax$. |
| 3a. Bei $\mathbf{K}aP, \mathbf{K}aX, \mathbf{K}bP, \mathbf{K}bX$ ist $a = b$ oder $P = X$. | 3b. Bei $\mathbf{K}aP, \mathbf{K}ax, \mathbf{K}bP, \mathbf{K}bx$ ist $a = b$ oder $P = x$. |
| 4a. Bei $\mathbf{K}aX, \mathbf{K}bX, \mathbf{K}cX$ gibt es $d = abc$. | 4b. Bei $\mathbf{K}ax, \mathbf{K}bx, \mathbf{K}cx$ gibt es $d = abc$. |
| 5a. Bei $\mathbf{K}aX, \mathbf{K}bX, abc = d$ ist $a = b$ oder $\mathbf{K}cX$. | 5b. Bei $\mathbf{K}ax, \mathbf{K}bx, abc = d$ ist $a = b$ oder $\mathbf{K}cx$. |
| 6a. Zu X gibt es $a \neq X$ ohne $\mathbf{K}aX$. | 6b. Zu x gibt es $a \neq x$ ohne $\mathbf{K}ax$. |

In der paarweisen Zuordnung der Axiome kommt die durchgängige Dualität zwischen Punkten und Geraden, die gemäß der Deutung der \mathbf{K} -Beziehung zugleich eine Dualität zwischen Inzidenz und Orthogonalität darstellt,

⁶⁾ Es würde übrigens genügen, einen Punkt vorauszusetzen.

zum Ausdruck. Bei Aufhebung der Beschränkung auf die durch x, y, X, Y bezeichneten Elemente erhält man durch Dualisierung eine Anzahl weiterer Axiomablen, deren wichtigste die Axiomablen 1c, 3c sind:

1c. (Vollleiptisches Axiom): Zu p, q gibt es A mit KAp, KAq .

3c. (Nichteuklidisches Axiom): Bei Kap, Kaq, Kbp, Kbq ist $a = b$ oder $p = q$.

Außer 1c, 3c sind unter den bei voller Heranziehung der Dualität hinzukommenden Axiomablen nur noch die folgenden *nicht beweisbar*:

1d. Zu P, Q gibt es A mit KAP, KAQ .

1e. Zu P, q gibt es A mit KAP, KAq .

1f. Zu p, q gibt es a mit Kap, Kaq .

2c. Bei KaB gibt es $C = aB$.

2d. Bei Kab gibt es $c = ab$.

4c. Bei KaP, KbP, KcP gibt es $D = abc$.

4d. Bei Kap, Kbp, Kcp gibt es $D = abc$.

4e. Bei Kap, Kbp, Kcp gibt es $d = Abc$

(und $\bar{d} = bAc$).

4f. Bei KAp, KBp, KcP gibt es $D = ABc$

(und $\bar{D} = AcB$).

4g. Bei KAp, KBp, KcP gibt es $d = ABC$.

Vollleipt.
Axiomablen.

5c. Bei $Kap, Kbp, abC = D$ ist $a = b$ oder KCP .

} Nichteuklid.

5d. Bei $KAp, KBp, ABC = D$ ist $A = B$ oder KCP .

} Axiomablen.

1d–f, 2c, d und 4c–g sind *vollleiptische Axiomablen*, d. h. äquivalent 1c. Man beweist nämlich zunächst leicht, daß die Forderung: „Es gibt $p = P$ “ mit der Forderung „Zu jedem p und ebenso zu jedem P gibt es $p = P$ “ äquivalent ist. Bei Gültigkeit der letzteren Formulierung gelten offenbar alle Axiomablen; die erste Formulierung folgt unmittelbar aus den Axiomen und jedem der Axiomablen 1d–f, 2c, d, 4c, d (1f wird man hier auf ein Paar p, q mit Kpq anwenden). Bei 4e–g hat man vorher mittels der Axiome 1a, b diejenigen Axiomablen zu beweisen, in denen D mit d vertauscht ist.

5c, d sind *nichteuklidische Axiomablen*, d. h. äquivalent 3c.

Beweis. 1. Es gelte Ax. 3c. Sei $Kap, Kbp, \alpha\beta = DC, \alpha \neq \beta$ (wobei α, β entweder für a, b oder für A, B stehen). Es gibt g mit Kcg, KDg (Ax. 1a), $\bar{c} = gC, \bar{d} = gD$ (Ax. 2a) (sowie gegebenenfalls $\bar{a} = Ap, \bar{b} = Bp$), $ab = DC = \bar{d}\bar{c}$ bzw. $\bar{a}\bar{b} = AB = DC = \bar{d}\bar{c}$. Axiom 5b lehrt Kag und Kbg (bzw. $K\bar{a}g$ und $K\bar{b}g$). Nach Ax. 3c ist $g = p$, also KCP .

7) Satz 4 (2) des § 3. — Zu den die Axiomablen betreffenden Beweisen, von denen hier nur einige angedeutet sind, reicht allgemein die Heranziehung weniger Sätze des § 3 (insbesondere 1 bis 5 und 40 (1)) aus.

— 2. Es gelte Axiomabel 5c bzw. d. Sei Kac , Kbc , Kad , Kbd , $a \neq b$. Nach Ax. 2b gibt es $\bar{A} = ac$, $\bar{B} = bc$, $\bar{C} = bd$, $\bar{D} = ad$. Man hat $\bar{A} \bar{B} \bar{C} = ab\bar{C} = ad = \bar{D}$. Daher ist nach Axiomabel 5c bzw. 5d: $K\bar{C}c$. Nun folgt nach Ax. 3a wegen $b \neq c$: $\bar{B} = \bar{C}$ und mithin $c = d$.

Alle übrigen Axiomablen, die sich bei der vollen Dualisierung ergeben, sind aus den Axiomen 1 bis 6a, b *beweisbar*; die Beweise seien hier unterdrückt, zumal sie (wie man sich etwa am Beispiel des Axiomabels „Bei KAp , Kbp , $Abc = d$ ist $A = b$ oder Kcp “ klarmacht) zum Teil ein wenig umständlich sind.

Wegen der vollen Dualität der vollelliptischen Geometrie lassen sich in ihr die Paare, die aus einem Punkt P und einer Geraden p mit $P = p$ bestehen, als neue Grunddinge (hier durch kleine gotische Buchstaben bezeichnet) einführen. Die Operation a stellt die Spiegelung an der a -Geraden, d. i. die Umwendung an dem zu ihr polaren a -Punkt dar. Kab bedeutet dann zugleich: 1. der a -Punkt liegt auf der b -Geraden, 2. der b -Punkt liegt auf der a -Geraden, 3. a - und b -Gerade stehen aufeinander senkrecht, 4. a - und b -Punkt sind konjugiert. Zusammen mit den trivialen Grundeigenschaften der Verknüpfung (S. 611, insbesondere der Streichungsregel) spannt das kurze *Axiomensystem*:

1. Zu a, b gibt es c mit Kac , Kbc .
2. Bei Kab gibt es $c = ab$.
3. Bei Kah , Kaf , Kbh , Kbf ist $a = b$ oder $h = f$.
4. Bei Kah , Kbh , Kch gibt es $d = abc$.
5. Bei Kah , Kbh , $abc = d$ ist $a = b$ oder Kch .
6. Zu a gibt es $b \neq a$ ohne Kab .

die gesamte vollelliptische Spiegelungsgeometrie auf.

Dem folgenden Paragraphen liegen die Axiome 1 bis 6a, b von S. 611 zugrunde.

§ 3.

Projektive Erweiterung.

1. (1) $aa \neq X$; $aa \neq x$. (2) Bei $X = ab$ ist Kab , KaX ; bei $x = ab$ ist Kab , Kax . (3) Bei Kab , KaX , KbX ist $X = ab$; bei Kab , Kax , Kbx ist $x = ab$.

Zum Beweise von (1): Die gegenteiligen Annahmen widersprechen den Ax. 6a bzw. b. — Zu (2): $a = b$ und ebenso $a = X$ bzw. x ist nach 1 (1) ausgeschlossen. — Zu (3): Für $P = ab$ (Ax. 2b) gilt nach Ax. 3a bzw. b: $P = X$ bzw. x .

2. (1) Zu a gibt es P mit KaP . (2) Bei Kab , Kcd , KaP , KbP , KcP KdP ist $ab = cd$.

(1) folgt aus Ax. 1b, 2b und 1 (2), (2) ist unmittelbare Folge von 1 (3).

3. Ergänzung zu den Axiomen 4a, b: (1) Bei $KaX, KbX, KcX, abc = d$ ist KdX . (2) Bei $Kax, Kbx, Kcx, abc = d$ ist Kdx .

Beweis. Bei $a = b$ ist $c = d$; bei $a \neq b$, $c = bad$ folgt die Behauptung nach Ax. 5a bzw. b.

4. (1) Bei KPQ gibt es $p = P$. (2) Bei Existenz eines $p = P$ gibt es zu jedem X ein $x = X$, zu jedem x ein $X = x$. (3) Bei Existenz eines $p = P$ gibt es zu jedem Paar q, r mit $q \neq r$ genau ein A mit KqA, KrA (vgl. Ax. 1c, 3).

Beweis zu (1): Es gibt a mit KaP, KaQ (Ax. 1a), $b = aP, p = aQ$ (Ax. 2a). Mit $bp = PQ$ ist Kbp , daher nach 1 (2), 1 (3): $p = P$. — Zu (2): Zu gegebenem X gibt es a mit KaX, KaP (Ax. 1a), also auch Kap . Es gibt weiter $b = aX, c = aP$ (Ax. 2a) = ap und wegen Kcp (1 (2)) auch $A = cp$ (Ax. 2b) = a . Da mit Kab (1 (2)) auch KAb ist, gibt es $x = bA$ (Ax. 2a) = $ba = X$. — (3) folgt unmittelbar aus (2) und Ax. 1a, 3a.

5. (1) Zu a, X gibt es $Y = aXa$; Abkürzung X_a . (2) Zu a, x gibt es $y = axa$; Abkürzung x_a . (3) Mit abc ist auch acb eine Gerade.

Man bedient sich zum Beweise von (1) eines h mit Kha, KhX (für $r = hX$ ist nach Ax. 4b ara Gerade), zum Beweise von (2) eines P und eines l mit KaP, KlP, Klx . — (3) ist unmittelbare Folge von (2). —

Die Sätze 43 bis 45, die das Ziel dieses Paragraphen darstellen, werden am kürzesten bewiesen, indem man sie einmal unter der Annahme, es gebe $p = P$, sodann unter der gegenteiligen Annahme herleitet. Im ersteren Falle lassen sich unmittelbare Beweise führen, im zweiten Falle werden die Sätze 6 bis 42 benutzt. Es ist daher zur Vereinfachung der Beweise zweckmäßig, den Sätzen 6 bis 42 die gemeinsame Prämisse, es gebe kein $p = P$, voranzustellen (doch sei angemerkt, daß der gesamte Beweisgang des § 3 sich unschwer auch ohne diese Annahme verfolgen läßt; es treten lediglich an verschiedenen Stellen Fallunterscheidungen hinzu).

Annahme A (für 6 bis 42): Kein Punkt ist einer Geraden gleich.

6. (Lotesatz.) Es sei $Kap, Kcq, par = \bar{a}, qcr = \bar{c}, Krs$. (1) Bei $b = psq$ gibt es $d = abc$, (2) umgekehrt ist bei $abc = d, a \neq c, KbT, KpT, KqT, KsT$, falls nicht zugleich KaT, KcT gilt, $b = psq$.

Beweis. (1) $K\bar{a}r, K\bar{c}r$. Nach Ax. 4b gibt es $h = \bar{a}\bar{s}\bar{c}$. $abc = appbc = apsqc = r\bar{a}s\bar{c}r = h_r$ (5 (2)). — (2) Da $\bar{a} = \bar{c}$ auf $pa = qc, p \neq q$ führt, würde nach Ax. 5a KaT, KcT sein; man hat also $\bar{a} \neq \bar{c}$. Gemäß Ax. 4a und 3 (1) gibt es $\bar{s} = pbq$ mit $K\bar{s}T$. $\bar{a}\bar{s}\bar{c} = r\bar{a}p\bar{p}bq\bar{c}r = d_r$. Daher ist nach 5 (3) auch $\bar{a}\bar{c}\bar{s}$ Gerade. Dann ist gemäß Ax. 5b $K\bar{s}r$ und mithin nach Ax. 3b $\bar{s} = s$.

Verabredungen zur Bezeichnung. Den folgenden Sätzen wird ein ausgezeichneter Punkt Z zugrunde gelegt. Ein Paar von Geraden a, p mit $a \neq p, KaZ$ lege ein „Hauptpaar“ $[ap]$ fest (die Schreibung $[ap]$ soll stets $a \neq p, KaZ$ in

sich schließen). Daß für eine Gerade q das Produkt paq eine Gerade ist, werde durch $q \subset [ap]$ bezeichnet. Zu gegebenem $X \neq Z$ ist durch $K h X$, $K h X$ das h eindeutig festgelegt; entsprechend sind zu gegebenem x stets durch $K h Z$, $K h x$, $F = h x$ sowohl h als F eindeutig (wegen Annahme \mathfrak{A}) festgelegt. Die konstruktive Abhängigkeit sei durch funktionale Kennzeichnung wiedergegeben: h wird als λX bzw. als λx bezeichnet, F als Φx . Bei $x = y$ ist $\Phi x = \Phi y$ und — falls nicht $\Phi x = Z$ — umgekehrt. Die für $X \neq Y$ durch $K r X$, $K r Y$ eindeutig festgelegte Verbindung r sei durch $\kappa(X, Y)$ bezeichnet, und für $\kappa(\Phi x, \Phi y)$ sei kurz $\kappa(x, y)$ geschrieben.

7. $a \subset [ap]$, $p \subset [ap]$.

8. Mit $q \subset [ap]$, $q \neq a$ ist $p \subset [aq]$.

9. Zu $[ap]$ und U gibt es ein q mit $K U q$, $q \subset [ap]$ und, falls nicht $K a U$, $K p U$ ist, auch nur ein solches q .

(Beweis mittels der Lotekonfiguration, 6.)

10. Bei $K a Z$, $K a S$, $K p S$, $a \neq p$ ist $q \subset [ap]$ mit $K q S$ gleichbedeutend. S heißt mit $[ap]$ *zusammengehörig*, $[ap]$ heißt *eigentlich*.

11. Mit einem Hauptpaar ist höchstens ein Punkt zusammengehörig.

12. Bei $K p Z$ gehört Z mit $[ap]$ zusammen, und umgekehrt.

Bei $K p(aZ)$ heißt $[ap]$ *ausgezeichnet*. Zwei Hauptpaare $[ap]$, $[bq]$ heißen *darstellungsgleich*, wenn entweder Z mit $[ap]$, $[bq]$ zusammengehört oder $a = b$, $q \subset [ap]$ ist (vgl. 8).

13. Für ein ausgezeichnetes $[ap]$ ist $q \subset [ap]$ mit $K q(aZ)$ gleichbedeutend. Zwei Hauptpaare $[ap]$, $[aq]$ sind also, falls ausgezeichnet, auch darstellungsgleich, und umgekehrt ist ein Hauptpaar, das einem ausgezeichneten Hauptpaar darstellungsgleich ist, ausgezeichnet.

Definition der Halbdrehung. Ein Geradenpaar u, v mit $uv \neq vu$, $K u Z$, $K v Z$ legt eine „Halbdrehung“ $\{uv\}$ fest (die Schreibung $\{uv\}$ soll stets die Voraussetzungen $K u Z$, $K v Z$, $uv \neq vu$ in sich schließen; da keine anderen Halbdrehungen benutzt werden, soll unter einer „Halbdrehung“ stets eine solche um Z verstanden werden). Sie sei dadurch charakterisiert, daß für irgendein x als „Bildgerade“ x^* die durch $K x^*(\Phi x)$, $K x^*(\lambda x \cdot uv)$ eindeutig bestimmte Gerade bezeichnet wird; — bei $K x Z$ ist also $x^* = xuv$.

14. Ist $x = y$, so auch $x^* = y^*$, und umgekehrt.

15. $(\lambda p)^* = \lambda(p^*)$; die Klammerung zum $*$ ist also bei λ unnötig.

16. Bei $K x Z$, $K y Z$ ist $xy = x^*y^*$.

17. Mit $[ap]$ ist $[a^*p^*]$ Hauptpaar, und umgekehrt.

18. Im speziellen Falle einer Halbdrehung $\{a \cdot \lambda p\}$ gehört Φp mit $[a^*p^*]$ zusammen.

19. Ist $[ap]$ ausgezeichnet, so auch $[a^*p^*]$, und umgekehrt.

20. Für jedes $[ap]$ folgt aus $K ap$: $K a^*p^*$, und umgekehrt. (Folge von 15.)

21. Mit $[ap]$ ist $[a^*p^*]$ eigentlich.

Beweis. $*$ beziehe sich auf die Halbdrehung $\{uv\}$. $[ap]$ gehöre mit P zusammen. Für $P = Z$ ist die Behauptung klar, für $uv = a \cdot \lambda p$ folgt sie aus 18. Andernfalls ist der durch $KPc, Ka^*c, Q = a^*c$ festgelegte (Ax. 1b, 2b) Punkt $Q \neq \Phi p$ (beide $\neq Z$). $\lambda \kappa(\Phi p, Q) = \lambda p \cdot aa^*$ (nach 6 (2)) $= \lambda p \cdot uv = \lambda p^*$ (15). Wegen $Kp^* \cdot \Phi p$ ist nach Axiom 3b $p^* = \kappa(\Phi p, Q)$, also außer dem nach 1 (2) gültigen Ka^*Q auch Kp^*Q .

22. Gehört P mit $[ap]$, Q mit $[a^*p^*]$ zusammen, so ist für beliebiges y KPy mit KQy^* äquivalent.

Zum Beweis. Daß KQy^* aus KPy folgt, ist für $y = a$ klar; für $y \neq a$ beweist es sich wie 21 (da dort Q unabhängig von p durch $KPc, Ka^*c, Q = a^*c$ bestimmt war). — Der Beweis der Umkehrung ist im Falle $P = Z$ einfach. Mit ausgezeichneten $[ap]$, $[a^*p^*]$ (19) gehört gemäß Annahme \mathfrak{A} kein Punkt zusammen. Andernfalls läßt sich für das w mit $KwP, Kw \cdot \lambda y$ (Ax. 1b) zeigen: $w^* = y^*$, und zwar bei $w = c$ unmittelbar (Ax. 3a, b), sonst ähnlich wie bei 21 an Hand von 6 (2).

23. $(\kappa(p, q))^* = \kappa(p^*, q^*)$ (nur für $\Phi p \neq \Phi q$ sinnvoll).

Zum Beweise wendet man 22 auf $\kappa(p, q)$ und die Hauptpaare $[\lambda p, p]$, $[\lambda q, q]$ an.

24. $x^{*o} = x^{oo}$.

Beweis. $*$ bzw. o mögen die Bilder bezüglich $\{uv\}$ und $\{st\}$ bezeichnen. Man darf $* \neq o$ und mithin $\Phi(x^*) \neq \Phi(x^o)$ annehmen. Für KxZ ist die Behauptung klar; andernfalls hat man nach 6 (2): $\lambda \kappa(x^*, x^o) = \lambda x^* \cdot \lambda x \cdot \lambda x^o = \lambda x^* \cdot st = \lambda x^{*o}$. Dann führt Ax. 3b auf $x^{*o} = \kappa(x^*, x^o)$. Entsprechend erhält man $x^{oo} = \kappa(x^*, x^o)$.

25. Mit $q \subset [ap]$ ist $q^* \subset [a^*p^*]$, und umgekehrt.

Beweis. Sei $q \subset [ap]$. Es kann $p \neq q$ angenommen werden. Nach 6 (2) ist $\lambda \kappa(p, q) = \lambda q \cdot a \cdot \lambda p$ und mithin nach 15 und 23: $\lambda \kappa(p^*, q^*) = \lambda q^* \cdot a^* \cdot \lambda p^*$ und nach 6 (1): $q^* \subset [a^*p^*]$. Die Schlußweise läßt sich umkehren.

26. (Freiheit der Punktbezeichnung.) Bei $q \subset [ap]$, $q, r \neq a$ ist $r \subset [ap]$ mit $r \subset [aq]$ äquivalent.

Nach 8 genügt es zu zeigen: Bei $q, r \subset [ap]$, $q \neq a$ ist $r \subset [aq]$.

Beweis. Für ausgezeichnetes $[ap]$ nach 13. Sonst sei die Halbdrehung $\{a \cdot \lambda p\}$ durch $*$ abgekürzt. $q^*, r^* \subset [a^*p^*]$ (25), also $Kq^* \cdot \Phi p, Kr^* \cdot \Phi p$ (18 und 10), $r^* \subset [a^*q^*]$ (14 und 10). Nun 25 Umkehrung.

27. Bei beliebiger Halbdrehung $*$ gibt es zu $[ap]$ ein $[bq]$ derart, daß $[b^*q^*]$ mit $[ap]$ darstellungsgleich ist.

Beweis. $*$ beziehe sich auf $\{cd\}$. Bei KpZ ist die Behauptung klar; bei Kap genügt das Hauptpaar $[(adc)p]$. — Sonst ist u mit $Ku \cdot \Phi p^*$,

$u \subset [ap]$ (9), $f = \lambda u \cdot dc$ (Ax. 4a) heranzuziehen. Es gibt q mit $Kq \cdot \Phi p$, $q \subset [u]$ (9); $f \neq u$ folgt aus dem Ausschluß der vorweggenommenen Fälle). Nach 18, 25 und 10 ist $Kq^* \cdot \Phi u$. Andererseits führt $q \subset [\lambda p, p]$ (10) nach 25 und 10 auf $Kq^* \cdot \Phi p^*$. Da der obengenannte Ausschluß auf $\Phi u \neq \Phi p^*$ führt, ist nach Ax. 3a $q^* = u$. — Nachdem man gegebenenfalls von p zu p_a übergegangen ist, kann man $cd \neq \lambda p \cdot a$ annehmen. (Dieser Übergang ist wegen der aus 26 folgenden Transitivität der Darstellungsgleichheit möglich.) Dann sind für $b = adc$ $[bq]$ und nach 17 $[b^*q^*]$ Hauptpaare. $[a^*q^*] = [au]$ ist darstellungsgleich $[ap]$ (8).

28. Sind bei $x \neq y$ xyu , xyv Geraden, so auch xuv , yuv .

Beweis. Der in den Sätzen 7 bis 27 zugrunde gelegte Punkt Z war keiner Bedingung unterworfen. Wird nun gemäß 2 (1) ein Z_0 mit KyZ_0 zugrunde gelegt, so liefert 26 (zusammen mit 5(3)) die Behauptung. (Im folgenden wird wieder ein beliebiges Z zugrunde gelegt.)

Definition des Idealpunktes.

1. Jeder Punkt P heiße ein *eigentlicher* Idealpunkt; KgP wird auch g/P geschrieben. 2. Bei gegebenen $g \neq h$, zu denen es keinen Punkt P mit g/P , h/P gibt, heiße die Gesamtheit Π aller Geraden u , für die ghu eine Gerade ist, ein *uneigentlicher* Idealpunkt; die Zugehörigkeit einer Geraden u zu Π wird durch u/Π mitgeteilt.

Wenn es zu dem (eentlichen oder uneentlichen) Idealpunkt Π ein f mit KfZ derart gibt, daß u/Π mit Kfu äquivalent ist, heiße Π *ausgezeichnet* über f ; Π ist dann gemäß der Annahme \mathfrak{A} (S. 614) uneentlich. — $[ap]$ soll den Idealpunkt Π „darstellen“, wenn a/Π , p/Π ist.

29. (1) Zu beliebigen Geraden g , h gibt es Π mit g/Π , h/Π . (2) Ein beliebiger Idealpunkt Π ist durch *jedes* Geradenpaar $g \neq h$ mit g/Π , h/Π eindeutig festgelegt; u/Π ist deduktiv gleichbedeutend damit, daß ghu Gerade ist.

Der Satz folgt aus Ax. 3a, 5a und 28.

30. (1) Zu Π , U gibt es q mit q/Π , q/U , (2) und zwar bei $\Pi \neq U$ genau eines.

Die Behauptung ergibt sich für eigentliches Π unmittelbar, sonst an Hand von 6 und 29(2).

31. (1) Zu jedem Idealpunkt Π gibt es Hauptpaare $[ap]$, die ihn darstellen. (2) Hauptpaare, die den gleichen Idealpunkt darstellen, sind darstellungsgleich, und umgekehrt.

Zum Beweise von (1) ist 30 (1) zu benutzen, zum Beweise von (2) entsprechend 30 (2) und 29 (2).

32. Das Hauptpaar $[ap]$ stelle den Idealpunkt Π dar. (1) Falls $[ap]$ ausgezeichnet ist, so auch Π , und umgekehrt; (2) falls P mit $[ap]$ zusammengehört, ist $P = \Pi$, und umgekehrt.

(1) folgt aus 13 und 29 (2), (2) aus Ax. 3a und 29 (2) allein.

Indem man bei gegebener Halbdrehung \ast und gegebenem Idealpunkt Π , der durch $[ap]$ dargestellt ist (31 (1)), den durch $[a^\ast p^\ast]$ (17) dargestellten Idealpunkt als das „Bild Π^\ast “ bezeichnet, hat man:

33. Zu jedem Idealpunkt Π gibt es bei gegebener Halbdrehung \ast (1) genau einen Bildpunkt Π^\ast , (2) genau einen Punkt Σ mit $\Sigma^\ast = \Pi$.

(1) folgt aus den Sätzen 17, 25 und 31, zum Beweise von (2) sind außerdem 27 und 14 heranzuziehen.

34. Zu einem uneigentlichen, jedoch nicht ausgezeichneten Π gibt es eine Halbdrehung \ast , so daß Π^\ast eigentlich ist (folgt aus 18, 31 (1) und 32).

35. Mit Π ist Π^\ast ausgezeichnet, und umgekehrt (nach 19, 31 (1) und 32 (1)).

36. Mit Π ist Π^\ast eigentlich (nach 21, 31 (1) und 32 (2)).

37. Aus g/Π folgt g^\ast/Π^\ast , und umgekehrt (25, 29 (2), 31 (1)).

38. Für irgendwelche Halbdrehungen \ast , \circ ist $\Pi^\ast \circ = \Pi \circ^\ast$ (24, 31).

39. Für einen Idealpunkt Π und eine Gerade g bestimmt die Gesamtheit der Geraden h , für die h_g/Π gilt, einen Idealpunkt Π_g (vgl. Def. zu 5 (2)).

Definition der Idealgeraden.

1. Eine Gerade heie eine *eigentliche* Idealgerade. 2. Eine Gesamtheit γ von lauter uneigentlichen Idealpunkten heie eine *uneigentliche* Idealgerade, wenn es eine eigentliche Gerade g und eine Z -Halbdrehung \ast derart gibt, da die Zugehörigkeit von Π zu γ mit g/Π^\ast gleichbedeutend ist. Die Zugehörigkeit von Π zu einer solchen uneigentlichen Idealgeraden γ wird mit γ/Π bezeichnet. 3. Die Gesamtheit ζ aller ausgezeichneten (und mithin uneigentlichen) Idealpunkte heie die *ausgezeichnete* (uneigentliche) Idealgerade; die Auszeichnung eines Π sei darum mit ζ/Π bezeichnet.

40. (1) Zu g gibt es $P \neq Q$ mit KgP, KgQ . (2) Zu einer Idealgeraden γ gibt es $\Pi \neq \Sigma$ mit $\gamma/\Pi, \gamma/\Sigma$.

Beweis zu (1). Für $l = gP$ mit KgP (2 (1) und Ax. 2a) beweist man an Hand der Axiome 1 bis 3 und 6 durch eine Fallunterscheidung, da es ein c ohne KcP und ohne Kcl gibt. Bei $KhP, Khc, C = hc, KkC, Kkg, Q = kg$ (Ax. 1b, 2b) ist nach 1 (2): KgQ ; mit Ax. 3a ergibt sich: $Q \neq P$. — (2) folgt für $\gamma \neq \zeta$ aus (1) und 33, für $\gamma = \zeta$ an Hand von 31 (2), 32 (1) und 40 (1).

41. Zu einer Idealgeraden γ und einer beliebigen Halbdrehung \ast gibt es (1) eine Idealgerade δ derart, da γ/Π mit δ/Π^\ast gleichbedeutend ist, (2) eine Idealgerade η , für die γ/Π^\ast mit η/Π gleichbedeutend ist.

δ wird als das „Bild γ^\ast “ von γ bezüglich der Halbdrehung \ast zu bezeichnen sein; nach 37 stimmt diese Definition für eine *eigentliche* Gerade $\gamma = g$ mit der früher gegebenen Definition (S. 615) überein. — Man hat dann als unmittelbare Folgerung von (1): 41 (3). Mit γ/Π ist γ^\ast/Π^\ast gleichbedeutend.

Beweis. 1. Annahme: $\gamma = \zeta$. Da ζ/Π mit ζ/Π^\ast gleichbedeutend ist (35), ist $\zeta = \zeta^\ast$. Damit sind (1), (2), (3) für ζ bewiesen.

2. Annahme: $\gamma = g$ eigentlich. (1) folgt aus 37. — Zu (2): Falls kein eigentlicher Punkt P mit g/P^* existiert, ist die Gesamtheit der Π mit g/Π^* nach Definition eine Idealgerade. Falls es aber einen eigentlichen Punkt P mit g/P^* gibt, sei $Q \neq P^*$ mit g/Q (nach 40 (1)) und $\Sigma^* = Q$, $P \neq \Sigma$ (33). Es gibt ein k mit k/P , k/Σ (30). Die Anwendung von (1) auf k führt mit Benutzung von Ax. 3a auf Beh. (2) für g .

3. Annahme: γ uneigentlich $\neq \zeta$, d. h. es gibt ein h und eine Halbdrehung \circ derart, daß γ/Ψ durch h/Ψ° definiert ist. Nach 40 (2) gibt es $\Pi \neq \Phi$ mit γ/Π , γ/Φ , d. h. h/Π° , h/Φ° , $\Pi^\circ \neq \Phi^\circ$ (33), wobei (wegen $\gamma \neq \zeta$) etwa Π als nicht ausgezeichnet angenommen werden darf. Schließlich sei Σ ein beliebiger Idealpunkt (also nicht notwendig $\neq \Pi, \Phi$) mit γ/Σ , d. h. h/Σ° . Zu (1): Es gibt eine Halbdrehung \bullet derart, daß Π^{**} eigentlich ist (34). Nach 30 gibt es ein k mit k/Π^{**} , k/Φ^{**} , also k°/Π^{***} , Φ^{***} (37). Wegen h/Π° , Φ° , Σ° ist $h^{**}/\Pi^{\circ**}$, $\Phi^{\circ**}$, $\Sigma^{\circ**}$ (37), h^{**}/Π^{***} , Φ^{***} , Σ^{***} (38). Da $\Pi^{***} \neq \Phi^{***}$ ist (33), ist $h^{**} = k^\circ$ (Ax. 3a) und somit k°/Σ^{***} . Dann ist nach 37 k/Σ^{**} . Nach dem unter der 2. Annahme geführten Absatz des Beweises gibt es ein — ebenso wie k unabhängig von dem variablen Σ bestimmtes — δ mit δ/Σ^* . Umgekehrt ergibt sich bei Rückwärtsdurchlaufung der Schlußkette für beliebiges Σ^* mit δ/Σ^* : γ/Σ . — Zu (2): Sei*) $A^* = \Pi$, $B^* = \Phi$, $\Gamma^* = \Sigma$ (33 (2)). Man hat $A \neq B$, A nicht ausgezeichnet (33, 35). Man schließt nun mit den gleichen Begründungen wie zu (1): Es gibt eine Halbdrehung \bullet derart, daß A^* eigentlich ist; es gibt u mit u/A^* , B^* , also $u^\circ/A^{\circ**}$, $B^{\circ**}$. Wegen $h/A^{\circ*}$, $B^{\circ*}$, $\Gamma^{\circ*}$ ist $h^\bullet/A^{\circ**}$, $B^{\circ**}$, $\Gamma^{\circ**}$. Da $A^{\circ**} \neq B^{\circ**}$, ist $h^\bullet = u^\circ$ und somit $u^\circ/\Gamma^{\circ**}$, u/Γ^* . Es gibt ein — wie u unabhängig von Σ bestimmtes — η mit η/Γ . Umgekehrt gilt für beliebiges Γ mit η/Γ : γ/Γ^* .

42. Zu p gibt es genau ein ausgezeichnetes Π mit p/Π .

Die Existenz eines solchen Π folgt aus 40 (1), die Eindeutigkeit an Hand von 1, 13, 29 (2), 31, 32.

43. (1) Zu $\alpha \neq \beta$ gibt es genau ein Π mit α/Π , β/Π . (2) Zu $\Pi \neq \Phi$ gibt es genau ein α mit α/Π , α/Φ .

Beweis. Zu (1): Bei $\alpha = \zeta$ folgt die Behauptung unmittelbar aus 42, 35, 33. Annahme: $\alpha, \beta \neq \zeta$. Es gibt die Halbdrehungen \bullet, \circ so, daß α^\bullet und β° und mithin $\alpha^{\circ\bullet}$ und $\beta^{\circ\bullet}$ eigentlich sind. Aus 41 (3) in Verbindung mit 38 (und bei eigentlichen α, β mit 40 (1)) folgt $\alpha^{\circ\bullet} \neq \beta^{\circ\bullet}$. Nach 29 gibt es daher genau einen Idealpunkt Σ mit $\alpha^{\circ\bullet}/\Sigma$, $\beta^{\circ\bullet}/\Sigma$. Es gibt genau ein Π mit $\Sigma = \Pi^{\circ\bullet}$ (33 (2)) = $\Pi^{\circ\bullet}$ (38), und da nach 41 (3) für irgendein Φ α/Φ , β/Φ mit $\alpha^{\circ\bullet}/\Phi^{\circ\bullet}$, $\beta^{\circ\bullet}/\Phi^{\circ\bullet}$ gleichbedeutend ist, hat man in Π den einzigen Punkt mit α/Π , β/Π .

*) Im folgenden sind die Buchstaben A, B, E, K, M, N, T stets, sofern nichts anderes vermerkt ist, als griechische Lettern, d. h. als Zeichen für Idealpunkte, zu lesen.

Zu (2): Daß nicht mehr als ein α mit α/Π , α/Φ existiert, folgt aus Abs. (1). Zum Nachweis der Existenz eines solchen α kann angenommen werden, daß nicht ζ/Π gelte. Falls nicht Π selbst eigentlich ist, wird eine Halbdrehung σ , für die Π^σ eigentlich ist, heranzuziehen sein (34). Es gibt ein h mit h/Π^σ , h/Φ^σ (30), also auch ein γ mit γ/Π , Φ (41).

Die Annahme \mathfrak{A} (S. 614), die bis hierhin als gemeinsame Prämisse galt, kann nun fallen. Zunächst ist 43 bei Existenz eines $p = P$ gültig, denn: (1) folgt dann wie oben aus 4 (3), und da nun alle Idealpunkte eigentlich sind, ist (2) durch Ax. 1a, 3a festgelegt. — Weiter werden für diesen Fall die Sätze 29, 30, 39 unmittelbare Folgen von 43 und den Axiomen. Andere Sätze werden in den nachfolgenden Beweisen von 44 und 45 nicht allgemein benötigt.

44. (Gegenpaarung.) $A, B, \Gamma, \Delta, E, \Phi$, seien die Paare konjugierter Ecken eines Viereits, dessen Seiten eigentlich sind und nicht sämtlich zusammenfallen⁹⁾. Es sei $A \neq \Gamma, \Delta$. Die sechs Ecken mögen von dem auf keiner Seite liegenden Idealpunkt Π aus durch die eigentlichen Geraden a, b, c, d, e, f angeschnitten werden. Dann ist mit $ac = db$ auch $ae = fb$.



Anmerkungen. 1. Unter Annahme eigentlicher Ecken, eines eigentlichen Aufpunktes und eines eigentlichen Schnittpunktes zweier durch

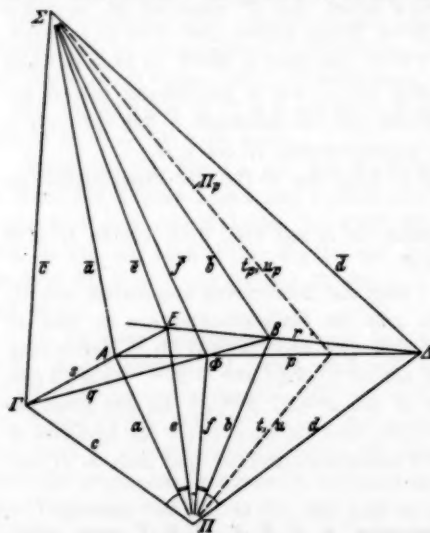


Fig. 1.

Produktbildung an Ecken entstehenden Geraden (insbesondere also unter Annahme \mathfrak{A}) wird der Beweis bereits hinter Satz 5 durchführbar. 2. Die Nebenbedingungen bezüglich der Lage von Π lassen sich abschwächen; doch genügt hier die vorliegende Fassung ($A \neq \Gamma, \Delta$ ist wesentlich).

Bezeichnungen und Zusammenstellung der vorausgesetzten Inzidenzen: $a, b, c, d, e, f/\Pi$, p, q, r, s nicht $/\Pi$, nicht $p = q = r = s$. $pas = \bar{a}$, $qbr = \bar{b}$, $qcs = \bar{c}$, $pdr = \bar{d}$, $res = \bar{e}$, $p/q = \bar{j}$; $ac = db$; $p, a, s, \bar{a}/A$, $q, b, r, \bar{b}/B$, $q, c, s, \bar{c}/\Gamma$, $p, d, r, \bar{d}/\Delta$, $r, e, s, \bar{e}/E$, $p, f, q, \bar{j}/\Phi$, $A \neq \Gamma, \Delta$.

Beweis. Mit $A \neq \Gamma$, s nicht $/\Pi$ ist $a \neq c$ (29(2) als Spezialfall von 43(2)). Es gibt $t = acf = dbf$ mit t/Π , $u = aed$ mit u/Π (29). $\bar{a}\bar{c}\bar{f} = pac/p = t_p$, $\bar{d}\bar{b}\bar{f} = pdb/p = t_p$, $\bar{a}\bar{e}\bar{d} = paed/p = u_p$, $t_p, u_p/\Pi_p$ (5(2) und 39). Sei $\bar{a}, \bar{c}/\Sigma$ (29). Aus $\bar{a} = \bar{c}$, d. h. $pa = qc$, und $a \neq c$ würde folgen p/Π . Also $\bar{a} \neq \bar{c}$. Entsprechend folgt $\bar{a} \neq \bar{d}$. Mit $\bar{a} \neq \bar{c}$ ist $\bar{f} \neq t_p$. Diese Ungleichungen führen zusammen mit den Produktdarstellungen von t_p, u_p — nach 29(2) und 5(3) — der Reihe nach auf $\bar{f}, t_p, \bar{d}, u_p/\Sigma$. Wäre $\Sigma = \Pi_p$, so hätte man — wegen $\Pi_p \neq A, \Delta, \Phi$ und $a_p, d_p, f_p/\Pi_p$ (39) — nach 29(2): $\bar{a} = a_p$, d. h. $p = s$, ebenso $\bar{d} = d_p$, d. h. $p = r$ und $\bar{f} = f_p$, d. h. $p = q$. Daher ist $\Sigma \neq \Pi_p$. Aus $t_p, u_p/\Pi_p$, Σ folgt dann nach 29(2): $t_p = u_p$, d. h. $ae = fb$.

45. (Pascalscher Satz.) Die Idealgeraden $\alpha, \beta, \gamma, \delta, \varepsilon, \vartheta$ mögen ein Sechseck bilden, dessen Idealecken abwechselnd auf zwei Idealgeraden liegen (wobei $\alpha \neq \delta, \beta \neq \varepsilon, \gamma \neq \vartheta$ sei). Die idealen Schnittpunkte von $\alpha, \delta, \beta, \varepsilon, \gamma, \vartheta$ liegen auf einer Idealgeraden.

Bezeichnung der Punkte und Zusammenstellung der Inzidenzen:

- | | | | |
|--------------------------------------|--------------------------------------|----------------------------------|------------------------------------|
| 1. $\alpha/A, \bar{A}, \Pi$, | 2. $\beta/\Gamma, \bar{A}, \Sigma$, | 3. $\gamma/\Gamma, \bar{A}, T$, | 4. $\delta/\Delta, \bar{A}, \Pi$, |
| 5. $\varepsilon/\Delta, K, \Sigma$, | 6. $\vartheta/A, K, T$, | 7. $\varphi/A, \Gamma, \Delta$, | 8. $\psi/\bar{A}, K, \bar{A}$. |

Behauptung: Es gibt $\omega/\Pi, \Sigma, T$.

Beweis. 1. Die acht vorausgesetzten Idealgeraden seien eigentlich und paarweise verschieden; die neun Idealpunkte seien ebenfalls paarweise verschieden, Π und K seien eigentlich. Π liegt dann auf keiner der Geraden $\beta, \gamma, \varepsilon, \vartheta, \varphi, \psi$. Sei $k/\Pi, K, g/\Pi, \Gamma, s/\Pi, \Sigma, t/\Pi, T$, (30), $n = gsk$ (Ax. 4a), $n, \varphi/N, r/K, N, \beta, r/A, \gamma, r/M, l/\Pi, A, m/\Pi, M$ (29, 30 bzw. 43). Aus $\Gamma = A$ oder $\Gamma = M$ würde $\varphi, r/A$ bzw. $/M$ und somit nach Ax. 3a folgen: $\Gamma = N$, also $g = n, s = k, s/\Pi$. Daher ist $\Gamma \neq A, M$, V. Läge Π auf r , so wäre, da φ nicht $/\Pi$ ist, $r = k = n, g = s, \beta/\Pi$. Doch ist r nicht $/\Pi$.

a) Man hat $gs = nk$. Das Vierseit $\beta s \varphi r$ hat die Eckenpaare $\Gamma, K, \Sigma, N, \Delta, A$. Es ist $\Gamma \neq \Sigma, N$. Satz 44 lehrt $g\delta = lk$ (unmittelbar dem dort angefügten Schema zu entnehmen).

b) Das Vierseit $\beta \gamma \psi r$ hat die Eckenpaare $\Gamma, K, \bar{A}, A, \bar{A}, M$. Es ist $\Gamma \neq \bar{A}, A$. Satz 44 lehrt $g\alpha = mk$.

c) Das Vierseit $\gamma \vartheta \varphi r$ hat die Eckenpaare Γ, K, A, M, T, N . Es ist $\Gamma \neq A, M$. Satz 44 lehrt $gt = nk$.

Die Gleichungen $gs = nk, gt = nk$ ergeben $s = t$.

2. Allein aus dem Inzidenzsatz 43 folgt ohne weiteres: a) Wenn in einer Konfiguration aus 9 Punkten und 9 Geraden (einschließlich ω) die acht ersten in der Zusammenstellung genannten Bedingungen erfüllt sind und außerdem zwei Punkte oder zwei Geraden übereinstimmen, so ist auch die neunte Bedingung erfüllt. b) Wenn in einer Konfiguration aus 9 Punkten und 9 Geraden

die Punkte und ebenso die Geraden paarweise verschieden sind, so sind alle Behauptungen, die die Abhängigkeit einer der neun Bedingungen von den acht übrigen fordern, äquivalent. Daher darf vorausgesetzt werden, daß in der dem Satze 45 zugrunde liegenden Konfiguration die Punkte paarweise verschieden, weiter die Geraden α bis $\varepsilon, \vartheta, \varphi, \psi$ paarweise verschieden und $\neq \zeta$ seien. Da, falls Π, Σ, T ausgezeichnet sind, die Behauptung gemäß der Definition von ζ erfüllt ist, kann außerdem angenommen werden, etwa Π sei nicht ausgezeichnet. Schließlich darf noch wegen $K \neq \bar{A}$ und $\psi \neq \zeta$ einer der Punkte K, \bar{A} , etwa K , als nicht ausgezeichnet vorausgesetzt werden.

Falls die Konfiguration nicht schon die Eigentlichkeitsvoraussetzung des Abs. 1 erfüllt, ist Annahme \mathfrak{A} in Kraft (vgl. S. 620); daher dürfen alle Hilfsätze benutzt werden. Nach 34, 41 (3) und der Definition der Idealgeraden gibt es dann eine endliche Folge von Halbdrehungen, die K und Π in eigentliche Punkte, α bis ε, ϑ und φ, ψ in eigentliche Geraden überführen. Die entstehende Figur erfüllt (nach 33, 41 (3), 43) die Voraussetzungen des Abs. 1; mit ihr erfüllt nach 41 (3) auch die ursprüngliche Figur die Behauptung.

Gemäß den Sätzen 43, 45 stellt der Bereich der Idealpunkte und -geraden eine projektive Erweiterung der Geometrie dar.

§ 4.

Die Polarkorrelation.

Den Begriff der Polarkorrelation in dem im vorigen Paragraphen erlangten Erweiterungsbereich der Spiegelungsgeometrie wird man axiomatisch in folgender Weise fassen können: Eine beiderseitig eindeutige Zuordnung Π der Geraden zu den Punkten hat Polarkorrelation zu heißen, wenn sie den Bedingungen genügt: 1. (involutorische Eigenschaft) bei Zuordnung $\Pi A\alpha, \Pi B\beta$, ist mit α/B auch β/A . 2. (Polareigenschaft) für zwei eigentliche Geraden a, b mit Kab zieht die Zuordnung ΠAa die Inzidenz b/A nach sich. Eine solche eindeutige Polarkorrelation setzt offenbar den nicht-euklidischen Fall, d. h. die Gültigkeit eines der nichteuklidischen Axiomen, voraus. Im Falle einer *linearen* Polarkorrelation gibt es bekanntlich einen Kegelschnitt, bezüglich dessen sie die im üblichen Sinne polare Zuordnung ist. Die Bewegungen der Geometrie stellen sich dabei als der „eigentliche“ Teil solcher Kollineationen dar, die den Kegelschnitt in sich überführen. Diesen Kegelschnitt hat man dem ins Einzelne gehenden Ausbau der Metrik zugrunde zu legen.

Es möge hier eine Bemerkung zur axiomatischen Einteilung der nicht-euklidischen Spiegelungsgeometrie in elliptische und hyperbolische eingeflochten werden. Für diese Einteilung pflegt man explizite auf den (nicht-zerfallenden) Kegelschnitt oder in anderer Weise auf die Polarkorrelation,

also jedenfalls auf den Erweiterungsbereich zurückzugreifen. Als ein ganz *eigentliches* und rein *figürliches* Axiom reicht nun (neben anderen) die folgende, verhältnismäßig übersichtliche Aussage aus:

Axiomabel H. Es gebe die verschiedenen (eigentlichen) Geraden a bis g und p, q mit folgenden Eigenschaften: pad, pae, qcf, qbg und bce sind Geraden, $K pa, K pb, K qc, K qd, K ef, ag = df$.

(Die Bedingungen, pad bis bce seien Geraden, lassen sich übrigens auch durch die schärfere Forderung eigentlicher Punkte P, A_1, B_1, C_1 mit $p, a, d, e/P, q, c, f/B_1, q, b, g/C_1, b, c, e/A_1$ ersetzen; — P und B_1 sind hier übrigens schon aus der ersten Fassung festgelegt.) **H** beschreibt eine Gegenpaarungskonfiguration mit dem Vierseit $A_1A_2B_1B_2\Gamma_1\Gamma_2$ und dem eigentlichen Aufpunkt P (Bezeichnungen der Figur). Daß bei Zugrundelegung eines der nicht-euklidischen Axiomablen das **H** notwendig und hinreichend für den Fall hyperbolischer Polarkorrelation ist, ergibt sich an Hand des Umstandes, daß der Gegenpaarungssatz, der zunächst für eigentliche Elemente bewiesen ist, im Erweiterungsbereich für jede Konfiguration mit eigentlichem Aufpunkt P gilt (die Gleich-

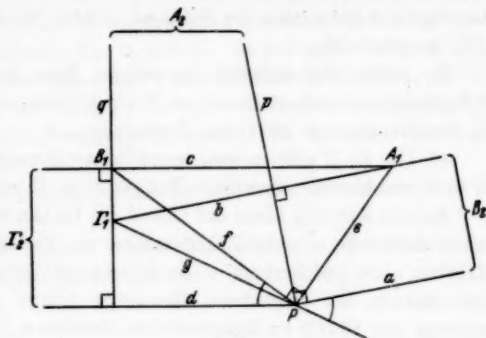


Fig. 2.

heit für Winkel an P bleibt ja bei Halbdrehung um P erhalten): 1. Bei Gültigkeit von **H** sind nach der evidenten Umkehrung des Gegenpaarungssatzes 44 A_2, B_2, Γ_2 kollinear. Da B_2 der Pol der Geraden p, Γ_2 der Pol von q ist, liegt A_2 auf seiner Polaren. 2. Liegt umgekehrt ein Idealpunkt A_2 auf seiner Polaren, so hat man bei der Konstruktion einer Konfiguration mit den in der Figur angegebenen Inzidenz-, Lot- und *Eigentlichkeitseigenschaften* nur etwa der Reihe nach mit q, A_1 (eigentlich), c, f, e, P, p zu beginnen.

Es sei hier nur kurz erwähnt, daß die Rolle des Axiomabels **H** übrigens nicht auf die Spiegelungsgeometrie beschränkt ist. In irgendwelchen Geometrien ohne Archimedisches Axiom bietet sich **H** als anordnungsloses Axiom dar, das die Forderung, die Winkelsumme im Dreieck sei $< \pi$ (die bekanntlich nicht der Forderung mehrerer Nichtschneidender zu einer Geraden durch einen außerhalb liegenden Punkt gleichkommt), zur Folge hat. — Bei Zulassung von $b = c, a = d, f = g$ hat man statt des hyperbolischen

ein „nichtelliptisches“ Axiomabel, das entsprechend die Forderung, die Winkelsumme sei $\leq \pi$, nach sich zieht.

Bei der Angabe einer linearen Polarkorrelation im Erweiterungsbereich der nichteuklidischen Spiegelungsgeometrie — sowie übrigens auch schon bei der entsprechenden Überlegung für den euklidischen Fall — ermöglicht die Gegenpaarung einen kurzen, synthetischen Weg. Der projektiven Erweiterung lagen ein ausgezeichnete Punkt Z und eine ausgezeichnete Idealgerade ζ , die seine Polare heißen möge, zugrunde. Die von Bachmann im Verfolg seiner Begründung der absoluten Geometrie (a. d. i. Anm. 3 (1) ang. O. S. 450) benutzten Gewebettranslationen gestatten, nun in evidenter Weise jedem eigentlichen Punkt eine Polare zuzuordnen. Dann läßt sich nach Bachmann-Reidemeister (a. d. i. Anm. 3 (2) ang. O. § 3) auf synthetischem Wege eine involutorische lineare Polarkorrelation angeben, sobald man nur bewiesen hat, daß die Involution der Schenkel rechter Winkel eines Büschels linear, d. h. projektiv ist.

Es bleibt also lediglich zu zeigen: Eine Rechtwinkelinvolution der Spiegelungsgeometrie stellt sich als Perspektivitätenkette dar. Zunächst steht in Erweiterung von 40 (1) zur Verfügung:

46. (1). Zu Q gibt es drei verschiedene Geraden $a, b, c/Q$. (2) Zu a gibt es drei verschiedene eigentliche Punkte P, Q, R mit $a/P, Q, R$.

Sodann läßt sich allein auf Grund des Inzidenzsatzes 43 und des Pascalschen Satzes 45 — ohne Zuhilfenahme von Punkten oder Geraden, deren Existenz nicht gesichert ist — der übliche synthetische Beweis für den Fundamentalsatz der projektiven Geometrie führen und weiterhin (mit Benutzung von 46 (2)) zu irgendwelchen Geraden $\alpha, \beta, \gamma, \delta$ eines Büschels mit $\gamma, \delta \neq \alpha, \beta$ eine Perspektivitätenkette angeben, die $\alpha, \beta, \gamma, \delta$ in $\beta, \alpha, \delta, \gamma$ überführt. Es gilt also in der Spiegelungsgeometrie der Satz:

47. Eine Perspektivitätenkette, die zwei verschiedene Geraden eines Büschels miteinander vertauscht, ist involutorisch.

48. Zu $g_1, g_2/P$ gibt es stets eine involutorische Perspektivitätenkette, die jede Gerade i/P mit $g_1 i g_2$ (insbesondere g_1 mit g_2) vertauscht.

Beweis. Offenbar läßt sich $g_1 \neq g_2$ voraussetzen, da man bei $g_1 = g_2$ nach 46 (1) von zwei verschiedenen Geraden $k, k_{g_1}/P$ ausgehen kann.

Außer g_1, g_2 gibt es $h_1 \neq g_1, g_2$ mit h_1/P (46 (1)). Sei $h_2 = g_1 h_1 g_2$, sei weiter i_1 irgendeine Gerade mit $i_1 \neq g_1, g_2$, i_1/P (i_1 braucht nicht notwendig von h_1, h_2 verschieden zu sein; diese Angabe soll den Fall berücksichtigen, in dem außer g_1, g_2, h_1, h_2 keine Gerade durch P unmittelbar als existent erwiesen ist); $i_2 = g_1 i_1 g_2$. Dann gibt es drei nichtkollineare eigentliche Punkte $U, V, W \neq P$ mit $i_1/U, g_1/V, g_2/W$. Ausgehend von diesen Punkten läßt sich auf Grund der Sätze 29, 30 und 44 unschwer ein Vierseit mit eigentlichen, nicht durch P gehenden und nicht sämtlich zusammen-

fallenden Seiten p, q, r, s konstruieren, dessen Eckenpaare, soweit sie bestimmt sind, auf $g_1, g_2, h_1, h_2, i_1, i_2$ liegen. Nun kann der übliche Beweis über Involutionen am Vierseit durchgeführt werden (p bis u Bezeichnungen der Figur):

$$g_1, g_2, h_1, i_1 \overset{p}{\wedge} t, g_2, s, q \overset{i_2}{\wedge} t, g_1, r, u \overset{q}{\wedge} g_2, g_1, h_2, i_2.$$

Die so fixierte, von der Wahl des i_1 unabhängige Perspektivitätenkette, die g_1 mit g_2 ($\neq g_1$) vertauscht und h_1 in h_2, i_1 in i_2 überführt, ist nach 47 involutorisch.

49. Für die Geraden $a_1, b_1, c_1, a_2, b_2, c_2$ eines Büschels sei $Ka_1a_2, Kb_1b_2, Kc_1c_2$. Es gibt dann eine involutorische Perspektivitätenkette, die a_1 mit a_2, b_1 mit b_2, c_1 mit c_2 vertauscht.

Beweis. Es darf $c_1 \neq a_1, a_2$ angenommen werden. Nach 48 gibt es nun eine Perspektivitätenkette, die a_1 mit c_1, a_2 mit $a_1a_2c_1 = c_2$ (2 (2)), b_1 mit $a_1b_1c_1$ vertauscht, und weiter eine Perspektivitätenkette, die a_2 mit c_1, a_1 mit $a_2a_1c_1 = c_2, b_2$ mit $a_2b_2c_1 = a_1b_1c_1$ vertauscht. Die durch Zusammensetzung entstehende Perspektivitätenkette, die a_1 in a_2, a_2 in a_1, b_1 in b_2, c_1 in c_2 überführt, ist nach 47 involutorisch.

Zusammen mit dem Fundamentalsatz der projektiven Geometrie führt 49 unmittelbar auf:

50. Die Schenkel der rechten Winkel eines Büschels gehen durch eine involutorische Perspektivitätenkette ineinander über.

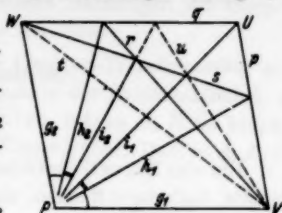


Fig. 3.

(Eingegangen am 18. 9. 1941.)

Eine Bemerkung zu den Plückerschen Formeln.

Von

Ott-Heinrich Keller in Berlin.

Hat eine Kurve C die Ordnung n , die Klasse k , das Geschlecht p , d Doppelpunkte, darunter r Spitzen, aber keine höheren Singularitäten, so gilt die bekannte Plückersche Formel:

$$(1) \quad k = n(n-1) - 2d - r^1)$$

und die Formel für das Geschlecht

$$(2) \quad 2p = (n-1)(n-2) - 2d.$$

Einer verwickelteren Singularität S kann man in bekannter Weise Äquivalenzzahlen zuordnen, die angeben, wie viele Doppelpunkte und wie viele Spitzen man sich in S vereinigt denken muß, damit (1) und (2) richtig bleiben. Diese Äquivalenzzahlen wurden jedoch, wie mir scheint, nirgends als wirkliche Anzahlen gedeutet. Nun scheint mir die Theorie der Nachbarpunkte, wie sie Enriques²⁾ entwickelt hat, eine Möglichkeit zu bieten, d und r in einfacher Weise zu deuten: Es ist in (1) und (2) $d = \sum \frac{i(i-1)}{2}$ zu setzen, wo i die Vielfachheiten aller mehrfachen (gewöhnlichen oder Nachbar-) Punkte durchläuft, und $r = \sum i$, der Summe der Vielfachheiten aller Satelliten.

Es sei mir gestattet, einen Beweis dieser Deutung mit den üblichen Methoden zu geben.

Die Anzahl der mehrfachen Nachbarpunkte sei a . Ist $a = 0$ und $r = 0$, so hat C in allen mehrfachen — etwa i -fachen — Punkten getrennte Tangenten. Die erste Polare und die Adjungierten sind dort jeweils $(i-1)$ -fach und schneiden C $i(i-1)$ -fach: dies ist also der Betrag, um den sich die Klasse und das doppelte Geschlecht durch das Hinzutreten des i -fachen Punktes erniedrigt haben, in Übereinstimmung mit der Behauptung. Wir können also einen Induktionsschluß nach a und r machen.

¹⁾ Der Übersichtlichkeit halber weiche ich von der üblichen Zählweise ab und zähle die Spitzen mit unter die Doppelpunkte.

²⁾ Enriques-Chisini, *Lezioni sulla teoria geometrica delle equazioni e delle funzioni algebriche*, Bd. 2.

O_0 sei ein singulärer i_0 -facher Punkt. Die Nachfolger ³⁾ von O_0 seien $O_{10}, O_{11}, \dots, O_{1\varrho_1}, O_{20}, \dots, O_{2\varrho_2}, \dots, O_{\alpha 0}, \dots, O_{\alpha\varrho_\alpha}$, und $O_{\lambda\mu}$ sei $i_{\lambda\mu}$ -fach. Dabei seien jeweils $O_{10}, O_{20}, \dots, O_{\alpha 0}$ unmittelbare Nachfolger, die übrigen Satelliten. A und B seien allgemeine Punkte der Ebene.

Wir üben nun auf C eine quadratische Cremona-Transformation mit den Hauptpunkten A, B, O_0 aus. Die Mittelpunkte der Strahlenbüschel, in die die Strahlenbüschel durch A, B, O_0 übergehen, mögen $\bar{A}, \bar{B}, \bar{O}_0$ heißen, \bar{C} geht in eine Kurve \bar{C} der Ordnung \bar{n} und der Klasse \bar{k} über. Da C in A, B, O_0 die Vielfachheiten $0, 0, i_0$ hatte, ist

$$(3) \quad \bar{n} = 2n - i_0 \text{ ⁴⁾ }.$$

\bar{C} hat in $\bar{A}, \bar{B}, \bar{O}_0$ die Vielfachheiten $(n - i_0), (n - i_0), n$ und geht durch diese Punkte längs linearer Zweige mit getrennten Tangenten.

Die Bilder der Nachfolger von O_0 liegen auf $\bar{A}\bar{B}$, und zwar die von $O_{10}, \dots, O_{\alpha 0}$ als gewöhnliche Punkte, die von $O_{\lambda\mu}$ mit $\mu \neq 0$ als Nachbarpunkte. \bar{C} geht durch sie mit gewissen Zweigen hindurch: die Summe ihrer charakteristischen Zahlen ist $\left(\sum_{\lambda=1}^{\alpha} i_{\lambda 0}, \sum_{\mu \neq 0} i_{\lambda\mu} \right)$. $\bar{A}\bar{B}$ rechnet also I -fach als Tangente von \bar{C} , wo $I = \sum_{\mu \neq 0} i_{\lambda\mu}$.

Wir müssen nun $\bar{k}, \bar{d}, \bar{r}, \bar{a}$ ausrechnen.

1. Um \bar{k} zu bestimmen, zählen wir die Tangenten von \bar{A} an \bar{C} ab. Hierzu gehören zunächst die k Bilder der k Tangenten von A an C . Dazu kommen die $(n - i_0)$ Tangenten an die Zweige durch \bar{A} , und jede zählt doppelt. Drittens haben wir die I -fache Tangente $\bar{A}\bar{B}$ zu beachten. Es ist also:

$$(4) \quad \bar{k} = k + 2(n - i_0) + I.$$

2. Um \bar{d} zu berechnen, haben wir zu beachten, daß von den mehrfachen Punkten von C einer, O_0 mit der Vielfachheit i_0 , weggefallen, dafür aber drei neue, $\bar{A}, \bar{B}, \bar{O}_0$ mit den Vielfachheiten $(n - i_0), (n - i_0), n$ hinzugetreten sind. Die übrigen mehrfachen Punkte sind unverändert übernommen worden.

$$(5) \quad 2\bar{d} = 2d - i_0(i_0 - 1) + 2(n - i_0)(n - i_0 - 1) + n(n - 1).$$

³⁾ Über die Begriffe „Nachbarpunkt“ und „Nachfolger“ siehe z. B. v. d. Waerden, Algebraische Geometrie, § 55. Berlin 1939.

⁴⁾ Eine Kurve vom Grad n , für die die Hauptpunkte einer quadratischen Transformation i_1, i_2, i_3 -fach sind, geht bekanntlich in eine Kurve vom Grad $2n - i_1 - i_2 - i_3$ und den Vielfachheiten $(n - i_2 - i_3), (n - i_1 - i_3), (n - i_1 - i_2)$ in den Hauptpunkten über.

3. Um \bar{r} zu berechnen, haben wir nur zu beachten, daß $\bar{O}_{i,\mu} (\mu \neq 0)$ aus der Liste der Satelliten zu streichen sind:

$$(6) \quad \bar{r} = r - \sum_{\substack{\mu=1 \\ \mu \neq 0}}^l i_{i,\mu} = r - I.$$

4. Es ist $\bar{a} = a - \sigma$, denn $O_{10}, O_{20}, \dots, O_{n,0}$ sind ja jetzt keine Nachbarpunkte mehr. Also sind \bar{a} und \bar{r} kleiner als a und r , und für \bar{C} gilt die Induktionsvoraussetzung

$$\bar{k} = \bar{n} (\bar{n} - 1) - 2\bar{d} - \bar{r}$$

und

$$2p = (\bar{n} - 1) (\bar{n} - 2) - 2\bar{d}.$$

Setzen wir hierin (3), (4), (5), (6) ein, so erhalten wir (1) und (2).

Die anderen Plückerschen Formeln folgen formal aus (1) und (2) und den dual entsprechenden.

(Eingegangen am 26. 5. 1942.)

Über den Ovalsatz von Böhmer-Mukhopadhyaya ¹⁾.

Von

Otto Haupt in Erlangen.

1. Einleitung.

1. 1. Der bekannte Böhmersche Ovalsatz²⁾ ist von S. Mukhopadhyaya³⁾ dahin verallgemeinert worden, daß irgend fünf Punkte eines beschränkten Ovals auf einer Ellipse liegen, falls alle sextaktischen (vgl. Nr. 2. 1.) Punkte des Ovals elliptisch sind; im übrigen stellt dabei Mukhopadhyaya nur die Forderung, daß das Oval differenzierbar sei, d. h. daß in jedem seiner Punkte genau eine Tangente existiere.

1. 2. Der Beweis von Mukhopadhyaya verläuft, kurz gesagt, so: Liegen fünf Punkte des Ovals \mathcal{C} auf einer Hyperbel \mathcal{H} , so liegen diese Punkte auf dem gleichen Hyperbelast \mathcal{H} und \mathcal{H} enthält sogar mindestens sechs Ovalpunkte $P_j, j = 1, \dots, 6$. Konstruiert man nun zu \mathcal{H} bzw. \mathcal{H} einen Kegelschnitt \mathcal{H}_1 , welcher mit \mathcal{C} fünf gewisse zwischen den P_j gelegene Punkte gemeinsam hat, so ist die numerische Exzentrizität^{3a)} von \mathcal{H}_1 größer als die von \mathcal{H} , also größer als 1, und folglich \mathcal{H}_1 ebenfalls eine Hyperbel. Ferner hat \mathcal{H}_1 mit \mathcal{C} sogar sechs, zwischen den P_j gelegene Punkte gemeinsam. Wiederholt man daher für \mathcal{H}_1 , statt für \mathcal{H} , diese Konstruktion usw., so ergibt sich eine Folge von Hyperbeln $\mathcal{H}_\nu, \nu = 1, 2, \dots$, derart, daß \mathcal{H}_ν mit \mathcal{C} mindestens sechs Punkte gemeinsam hat, welche für hinreichend große ν in beliebiger Nähe eines Punktes Q von \mathcal{C} liegen. Dann ist Q sextaktischer, aber nicht elliptischer Punkt von \mathcal{C} im Widerspruch mit der Voraussetzung.

¹⁾ Vgl. eine Voranzeige in Haupt, Bemerkung über parabolisch konvexe und konkave Ovale. Sitz.-Ber. d. physikal.-medizin. Sozietät Erlangen 72 (1940/41), S. 216ff., Nr. 5. 1. Vgl. ferner die ebendort (Nr. 6) angezeigten Ergebnisse von Herrn Hjelmalev.

²⁾ P. Böhmer, Über elliptisch-konvexe Ovale. Math. Annalen 60 (1905), S. 256ff.

³⁾ S. Mukhopadhyaya, Generalized form of Böhmer's theorem for an elliptically curled non-analytic oval. Math. Zeitschr. 30 (1929), S. 560ff., sowie Collected geometrical papers, Part I, Calcutta (1929), S. 33ff.

^{3a)} Unter der numerischen Exzentrizität von $a^{-2}x^2 \pm b^{-2}y^2 - 1 = 0$ verstehen wir die Zahl $e = (1 \mp b^2a^{-2})^{\frac{1}{2}} \geq 0$, wobei für den Fall der Ellipse $0 < b \leq a$ sein soll. Für $e > 1$ gehört zum größeren e der größere Asymptotenwinkel; dabei verstehen wir unter dem Asymptotenwinkel einer Hyperbel \mathcal{H} das Winkelmaß derjenigen beiden, von den Asymptoten der \mathcal{H} gebildeten Scheitelwinkelräume, innerhalb deren die beiden Äste von \mathcal{H} liegen; Asymptotenwinkel und numerische Exzentrizität bestimmen einander ein-eindeutig. Für die Parabel ist $e = 1$.

1. 3. Bei näherer Betrachtung des in Nr. 1. 2. angedeuteten Beweises zeigt sich nun: Die erwähnte Konstruktion, von welcher in abzählbar vielen diskreten Schritten die Folge der \mathfrak{S} , geliefert wird, läßt sich ersetzen durch eine stetige Abänderung von \mathfrak{S} , welche durch eine gewisse stetige, monotone Zusammenziehung der P , auf einen Punkt Q von \mathfrak{C} bestimmt wird. Bei einer solchen Abänderung nimmt wieder, wie leicht zu sehen, der Asymptotenwinkel (und folglich^{3a)} die numerische Exzentrizität] von \mathfrak{S} im wesentlichen nicht ab. Bei dieser Wendung des Beweises in einen Kontinuitätsbeweis⁴⁾ wird die Differenzierbarkeit von \mathfrak{C} (vgl. Nr. 1. 1.) nicht mehr benötigt; auch scheint uns der Kontinuitätsbeweis durchsichtiger zu sein. Zugleich aber führt der Kontinuitätsbeweis unmittelbar zur folgenden *Verallgemeinerung* des Böhmer-Mukhopadhyayaschen Satzes: Besitzt jeder sextaktische Kegelschnitt eines Ovals \mathfrak{C} (d. h. jeder in einem Ovalpunkt mindestens „sechspunktig berührende“ Kegelschnitt) eine numerische Exzentrizität, die kleiner ist als $\epsilon \geq 1$, so ist auch die Exzentrizität eines jeden Kegelschnittes durch fünf beliebige Punkte von \mathfrak{C} kleiner als ϵ (vgl. die genauere Formulierung in Nr. 2. 1. 1.).

1. 4. Der in Nr. 1. 3. besprochene Kontinuitätsbeweis gilt allgemeiner auch dann, wenn das System der Kegelschnitte ersetzt wird, beispielsweise durch eine System von (nicht notwendig algebraischen) Kurven zweiter Ordnung, deren jede durch $(2k + 1)$ Punkte des Ovals eindeutig bestimmt ist und die paarweise höchstens $2k$ Punkte gemeinsam haben ($2 \leq k$). Da indes im wesentlichen alles schon am Falle der Kegelschnitte deutlich wird, beschränken wir uns in vorliegender Mitteilung auf diesen (speziellen) Fall und verweisen hinsichtlich der erwähnten allgemeineren Fälle auf Ausführungen an anderer Stelle⁵⁾.

Jedenfalls erweist sich so der Satz von Böhmer-Mukhopadhyaya als Spezialfall eines Theorems über geometrische Ordnungen, m. a. W. als ein Satz von stark topologischem Gehalt.

^{3a)} Über Kontinuitätsbetrachtungen in anderen Arbeiten von Mukhopadhyaya vgl. auch die Hinweise in Haupt, Zur Theorie der Realitätsordnungen. Monatsh. f. Math. u. Phys. 40 (1933), S. 1ff., Nr. 0. 4. — Bei der Konstruktion von Mukhopadhyaya müssen Punkte von $\mathfrak{C} \mathfrak{S}$, unter Umständen mehrfach gezählt werden. Für eine Definition der entsprechenden Vielfachheitszählung ist aber die Forderung der Differenzierbarkeit von \mathfrak{C} kaum zu entbehren. Beim Kontinuitätsbeweis hingegen können o. B. d. A. die Punkte von $\mathfrak{C} \mathfrak{S}$ stets als voneinander verschiedene Schnittpunkte angenommen werden, womit sich eine besondere Vielfachheitszählung und damit eine Annahme über die Differenzierbarkeit von \mathfrak{C} erübrigt.

⁵⁾ Vgl. Haupt, Über Verallgemeinerungen des Böhmerschen und verwandter Ovalsätze, Abh. math. Semin. Hamburg, Univ. 1943; ferner Über einige affingometrische Ovalsätze in der direkten Infinitesimalgeometrie (Erscheint an anderer Stelle).

2. Formulierung und Beweis für den Fall der Kegelschnitte.

2.1. Im folgenden werde als *Oval* \mathfrak{C} jedes topologische, *keine Strecken enthaltende* Kreisbild in der affinen Ebene bezeichnet, welches die lineare Ordnung Zwei besitzt. Insbesondere ist also jedes Oval \mathfrak{C} *beschränkt* sowie *konvex* im üblichen Sinne.

Ein Punkt S des Ovals \mathfrak{C} heiße von der Kegelschnittordnung Sechs, kürzer: *sextaktisch*, wenn in beliebig kleiner Umgebung von S auf \mathfrak{C} (mindestens) sechs (verschiedene) Punkte existieren, welche auf einem Kegelschnitt⁶⁾ liegen⁷⁾. Ferner sagen wir, ein Punkt P von \mathfrak{C} besitze die obere bzw. untere Exzentrizitätsschranke $e \geq 1$, kürzer: Es sei P ein ($< e$)-Punkt bzw. ein ($\geq e$)-Punkt von \mathfrak{C} , wenn folgendes gilt: Irgend fünf, zu P hinreichend benachbarte (verschiedene) Punkte von \mathfrak{C} liegen auf einem Kegelschnitt, dessen numerische Exzentrizität kleiner bzw. nicht kleiner ist als e . Die (< 1)-Punkte heißen speziell auch *elliptische Punkte*⁸⁾.

2.1.1. Der auf den Fall der Kegelschnitte bezügliche, nachstehend zu beweisende Satz lautet:

Satz. Vor. *Es sei $e \geq 1$ vorgegeben. Ferner seien die sextaktischen Punkte des vorgelegten Ovals \mathfrak{C} sämtlich ($< e$)-Punkte.*

Beh. *Je fünf Punkte von \mathfrak{C} liegen auf einem Kegelschnitt, dessen numerische Exzentrizität kleiner ist als e .*

Zum Beweis zeigen wir: Keine fünf Punkte von \mathfrak{C} liegen auf einem Kegelschnitt, dessen Exzentrizität mindestens gleich e ist.

2.2. Zu Abkürzung bezeichnen wir beim Beweis jeden (nicht ausgearteten) Kegelschnitt, dessen numerische Exzentrizität kleiner bzw. nicht

⁶⁾ Ein Kegelschnitt, welcher mit einem Oval mindestens 5 verschiedene Punkte gemeinsam hat, kann *nicht ausgeartet* sein; denn andernfalls müßten mindestens 3 Punkte des Ovals auf einer Geraden liegen, während doch ein Oval hier als streckenfrei vorausgesetzt wird.

⁷⁾ Jedes Oval besitzt sextaktische Punkte, wie schon aus dem oben im Text nachfolgenden Kontinuitätsbeweis hervorgeht. — Vgl. auch S. Mukhopadhyaya, *New methods in the geometry of a plane arc*, I. *Cyclic and sextactic points*. Bull. Calcutta Math. Soc., vol. 1 (1909) = *Collected geometrical papers*, Part I, Calcutta 1929, S. 13ff., wo die Mindestanzahl sextaktischer Punkte eines Ovals bestimmt wird. Vgl. dazu noch S. Mukhopadhyaya, *Extended minimum-number theorems of cyclic and sextactic points on a plane convex oval*. Math. Zeitschr. 33 (1931), S. 648ff. = *Collected geometrical papers*, Part II, Calcutta 1931, S. 159ff. In beiden Arbeiten werden Differenzierbarkeitsannahmen gemacht.

⁸⁾ Betr. elliptische und sextaktische Punkte in der affinen Differentialgeometrie vgl. W. Blaschke, *Vorlesungen über Differentialgeometrie*, II. *Affine Differentialgeometrie*, bearbeitet von K. Reidemeister, Berlin 1923, S. 27/28, sowie S. 43ff. Betr. den Böhmerschen Satz vgl. S. 47ff.

kleiner bzw. größer ist als e als einen $(< e)$ - bzw. $(\geq e)$ - bzw. $(> e)$ -Kegelschnitt, kürzer: $(< e)$ -KS bzw. $(\geq e)$ -KS bzw. $(> e)$ -KS.

2.2.1. Aus der Voraussetzung des behaupteten Satzes (Nr. 2.1.1.) folgt zunächst, daß *jeder* $(\geq e)$ -KS mit \mathfrak{C} *nur endlich viele Punkte gemeinsam hat*. Andernfalls nämlich enthielte der Durchschnitt dieses Kegelschnittes mit \mathfrak{C} (mindestens) einen Häufungspunkt H auf \mathfrak{C} und H wäre jedenfalls sextaktischer Punkt und zugleich $(\geq e)$ -Punkt.

2.2.2. Wir bemerken zu Nr. 2.2.1. noch:

Zu jedem $(\geq e)$ -KS \mathfrak{K} existiert immer ein $(> e)$ -KS \mathfrak{K}' , welcher folgende Eigenschaft besitzt: Es ist \mathfrak{K}' zu \mathfrak{K} beliebig benachbart; ferner hat \mathfrak{K}' mit \mathfrak{C} nicht weniger Punkte gemeinsam als \mathfrak{K} und die Punkte von $\mathfrak{K}'\mathfrak{C}$ sind sämtlich *Schnittpunkte*.

In der Tat: 1. Wegen $1 \leq e$ besitzt \mathfrak{K} zwei (eventuell zusammenfallende) reelle uneigentliche Punkte U' , U'' . Diejenigen Kegelschnitte $\bar{\mathfrak{K}}'$, welche U' und U'' enthalten, sind *Erstens* durch Vorgabe von drei eigentlichen Punkten eindeutig bestimmt und *Zweitens* sind sie sämtlich $(\geq e)$ -KS. Wegen *Erstens* und Nr. 2.2.1. gibt es unter den $\bar{\mathfrak{K}}'$ solche \mathfrak{K}' , welche zu \mathfrak{K} beliebig benachbart sind, für welche ferner $\mathfrak{K}'\mathfrak{C}$ (endlich viele und) mindestens ebenso viele Punkte enthält wie $\mathfrak{K}\mathfrak{C}$ und wobei alle Punkte von $\mathfrak{K}'\mathfrak{C}$ Schnittpunkte sind⁹⁾. — 2. Ist etwa e'' die Exzentrizität von \mathfrak{K}' , so gibt es zu \mathfrak{K}' beliebig benachbarte Kegelschnitte $\bar{\mathfrak{K}}'$, deren Exzentrizität größer ist als e'' und mit welchen \mathfrak{C} mindestens ebenso viele Punkte gemeinsam hat wie mit \mathfrak{K}' , wobei überdies sämtliche Punkte von $\bar{\mathfrak{K}}'\mathfrak{C}$ Schnittpunkte sind. Man gelangt nämlich bei Festhalten etwa von U' sowie von drei eigentlichen Punkten von \mathfrak{K}' zu Kegelschnitten $\bar{\mathfrak{K}}$ größerer Exzentrizität, sobald man nur den U'' beliebig wenig, aber passend ändert. Da $\mathfrak{K}'\mathfrak{C}$ lauter Schnittpunkte, und zwar nur endlich viele, enthält, so hat jeder, zu \mathfrak{K}' hinreichend benachbarte Kegelschnitt $\bar{\mathfrak{K}}$ mindestens ebenso viele Schnittpunkte mit \mathfrak{C} gemeinsam wie \mathfrak{K}' ; und von $\bar{\mathfrak{K}}$ gelangt man wie in 1. zu einem zu $\bar{\mathfrak{K}}$ beliebig benachbarten Kegelschnitt \mathfrak{K}' mit größerer Exzentrizität als e'' , für welchen $\mathfrak{K}'\mathfrak{C}$ nur Schnittpunkte, und zwar nur endlich viele, enthält.

2.3. Die Behauptung in Nr. 2.1.1. wird jetzt indirekt bewiesen. Wir nehmen also die Existenz von (mindestens fünf) Punkten P_j , $j = 1, \dots, k$; $5 \leq k$, auf \mathfrak{C} an, welche Schnittpunkte eines $(\geq e)$ KS mit \mathfrak{C} sind. Gemäß Nr. 2.2.2. kann dabei o. B. d. A. angenommen werden, daß dieser $(\geq e)$ -KS ein $(> e)$ -KS \mathfrak{K} ist und daß $\mathfrak{K}\mathfrak{C}$ nur Schnittpunkte enthält. Wegen $e \geq 1$

⁹⁾ Vgl. Haupt, a. a. O. ⁴⁾, Nr. 1.6., a) und Nr. 2.4.

¹⁰⁾ Vgl. Mukhopadhyaya, a. a. O. ³⁾, Lemma I.

ist \mathfrak{H} eine Hyperbel. Wegen $5 \leq k$ liegen¹⁰⁾ die P_j und überhaupt alle Punkte von $\mathfrak{H}\mathfrak{C}$ auf dem gleichen Aste \mathfrak{A} von \mathfrak{H} . Da \mathfrak{C} eine geschlossene Kurve ist, muß $k \geq 6$ sein. Da ferner \mathfrak{C} sowohl als \mathfrak{A} konvex sind, liegen¹¹⁾ die P_j simultan-natürlich geordnet auf \mathfrak{C} und auf \mathfrak{A} , d. h. es existiert eine Anordnung der P_j , etwa P_1, \dots, P_k derart, daß die P_j in dieser Anordnung (Reihenfolge) sowohl auf \mathfrak{C} als auf \mathfrak{A} aufeinanderfolgen, wenn \mathfrak{C} bzw. \mathfrak{A} in geeigneter Richtung durchlaufen (orientiert) werden. Da ferner \mathfrak{A} durch fünf seiner Punkte eindeutig bestimmt ist und sich stetig mit diesen Bestimmungspunkten ändert, kann man auf irgend sechs, unmittelbar aufeinanderfolgende der P_j den Kontraktionssatz¹²⁾ anwenden. Infolgedessen lassen sich diese sechs Punkte stetig und monoton in beliebige Nähe eines gewissen Punktes Q von \mathfrak{C} zusammenziehen derart, daß die sechs Punkte in jedem Augenblick des Zusammenziehungsprozesses, kurz: der Kontraktion, auf einem Kegelschnitt \mathfrak{K} liegen. Überdies kann kein im Laufe eines solchen Prozesses auftretender Kegelschnitt ausgeartet sein⁶⁾. Es genügt daher, folgendes zu zeigen:

Beh. A. Bei passender Wahl derjenigen 6 unter den P_j , mit welchen man die Kontraktion beginnt, ist der Kegelschnitt \mathfrak{K} in jedem Augenblick der Kontraktion ein ($> e$)-KS, so daß also insbesondere, wegen $1 \leq e$, die sechs Punkte immer auf dem gleichen Aste einer nicht-ausgearteten Hyperbel liegen.

Aus der Beh. A. folgt nämlich, daß Q zwar sextaktisch, aber nicht ($< e$)-Punkt ist, womit die Beweisannahme ad absurdum geführt wird.

2.4. Beim Beweise der Beh. A. (vgl. Nr. 2.3.) bedienen wir uns folgender Bezeichnung: Es seien Q_1, \dots, Q_6 irgend sechs unmittelbar auf \mathfrak{C} und auf dem gegebenen Hyperbelast \mathfrak{A} aufeinanderfolgende Punkte unter den Schnittpunkten P_j von \mathfrak{A} und \mathfrak{C} ; weiter sei \mathfrak{A}' der von Q_1 und Q_6 begrenzte, abgeschlossene, alle Q_1, \dots, Q_6 enthaltende Teilbogen von \mathfrak{A} und \mathfrak{A}'' sei das Komplement von \mathfrak{A}' bezüglich \mathfrak{A} . Das 6-tupel Q_1, \dots, Q_6 werde als ein äußerliches bezüglich \mathfrak{C} dann bezeichnet, wenn eine Umgebung von Q_1 auf \mathfrak{A}'' im Äußeren $a(\mathfrak{C})$ von \mathfrak{C} liegt¹³⁾; da die Q_j sämtlich Schnittpunkte von \mathfrak{A} mit \mathfrak{C} sein sollen, liegt dann eine Umgebung von Q_1 auf \mathfrak{A}' im Inneren $i(\mathfrak{C})$ von \mathfrak{C} , ferner eine Umgebung von Q_6 auf \mathfrak{A}' bzw. auf \mathfrak{A}'' im Inneren bzw. im Äußeren von \mathfrak{C} .

¹¹⁾ Vgl. z. B. Haupt, Bestimmung der zyklisch ordnungshomogenen ebenen Bogen, 1. Mitt. Journ. f. d. r. u. angew. Math. 178 (1937), Nr. 2. 3.

¹²⁾ Haupt, a. a. O. ⁴⁾, Nr. 4. 4. ff. Hinsichtlich einer arithmetischen Fassung vgl. Haupt, Bestimmung der zyklisch ordnungshomogenen ebenen Bogen, 2. Mitt. Journ. f. d. r. u. angew. Math. 180 (1939), S. 54, Zusatz.

¹³⁾ Unter dem Äußeren $a(\mathfrak{A})$ oder $a(\mathfrak{C})$ bzw. dem Inneren $i(\mathfrak{A})$ oder $i(\mathfrak{C})$ von \mathfrak{A} oder \mathfrak{C} verstehe man den offenen Kern des Komplementes der konvexen Hülle bzw. den der konvexen Hülle von \mathfrak{A} oder \mathfrak{C} . [Komplement bezüglich der (affinen) Ebene.]

Weil der (zu Beginn der Kontraktion vorliegende) Hyperbelast \mathfrak{H} beiderseits unbeschränkt ist, läßt sich aus den Punkten P_j von $\mathfrak{H}\mathfrak{C}$ immer (mindestens) ein äußerliches 6-tupel herausgreifen. Es sei ein solches 6-tupel gewählt und dieses werde gebildet aus den in dieser Reihenfolge auf \mathfrak{H} unmittelbar aufeinanderfolgenden (Schnitt-) Punkten Q_1, \dots, Q_6 . Dieses 6-tupel ist nun einer Kontraktionsfolge im Sinne des Kontraktionssatzes¹²⁾ zu unterwerfen. Eine solche Kontraktionsfolge setzt sich zusammen aus Operationen von dreierlei Art: (I) werden vier von den Q_i festgehalten und die übrigen beiden auf \mathfrak{C} gegeneinander so bewegt, daß der kleinste, die sechs Punkte enthaltende, also von Q_1 und Q_6 begrenzte, Teilbogen von \mathfrak{C} dabei nicht zunimmt; sobald Stützpunkte zwischen Q_1 und Q_6 auftreten, ist die Operation jedenfalls abzubrechen. — (II) werden die jeweils vorliegenden sechs Punkte einer geeigneten, aber beliebig klein wählbaren Verrückung auf \mathfrak{C} unterworfen. — (III) Falls zwischen den Punkten des jeweils vorliegenden 6-tupels Gewinne stattfinden, wird aus den alsdann vorhandenen (mehr als sechs) Punkten ein anderes, ebenfalls bezüglich \mathfrak{C} äußerliches 6-tupel ausgewählt und dieses der anschließenden Operation zugrunde gelegt.

2.4.1. Wir haben jetzt zu zeigen, daß bei den soeben aufgezählten Operationen (I) bis (III) immer nur ($> c$)-KS erhalten werden, vorausgesetzt, daß die Änderungen bei (II) sämtlich hinreichend klein gewählt werden.

2.4.1.0. Zunächst bemerken wir:

Es sei Q_1, \dots, Q_6 ein äußerliches 6-tupel bezüglich \mathfrak{C} . Ferner bezeichne \mathfrak{C}_r bzw. \mathfrak{H}_r denjenigen offenen, von Q_r und Q_{r+1} begrenzten Teilbogen von \mathfrak{C} bzw. von \mathfrak{H} , welcher keine Q_i enthält, $r = 1, \dots, 5$; $\varrho = 1, \dots, 6$. Dann liegt $\mathfrak{C}_{2\mu-1}$ im Äußern $a(\mathfrak{H})$ von \mathfrak{H} und folglich $\mathfrak{C}_{2\sigma}$ im Innern $i(\mathfrak{H})$ von \mathfrak{H} , $\mu = 1, 2, 3$; $\sigma = 1, 2$.

Bew. Weil $\mathfrak{H}\mathfrak{C}$ außer Q_1 und Q_2 noch weitere Punkte enthält, liegen \mathfrak{H}_1 und \mathfrak{C}_1 auf der gleichen Seite der Verbindungsgeraden von Q_1 und Q_2 ; denn andernfalls würde wegen der Konvexität von \mathfrak{H} und \mathfrak{C} sowie wegen $\mathfrak{H}_1\mathfrak{C} = \mathfrak{C}_1\mathfrak{H} = 0$ sein: $\mathfrak{H}\mathfrak{C} = (Q_1) + (Q_2)$. Da aber nach Voraussetzung $\mathfrak{H}_1 \subset i(\mathfrak{C})$ gilt, so folgt $\mathfrak{C}_1 \subset a(\mathfrak{H})$, also auch (weil alle Q_i Schnittpunkte sind) $\mathfrak{C}_{2\mu-1} \subset a(\mathfrak{H})$; w. z. b. w.

2.4.1.1. *Betr. die Operationen* (I). Es sei Q_r der, von Q_1 an gerechnet, erste unter den Q_1, \dots, Q_6 , welcher bei der betrachteten Operation (I) beweglich ist ($1 \leq r \leq 5$). Falls $r \equiv 1 \pmod{2}$ ist, liegt \mathfrak{C}_r in $a(\mathfrak{H})$ (vgl. Nr. 2.4.1.0). Und da Q_r bei einer Operation (I) (Kontraktion) auf \mathfrak{C} gegen Q_{r+1} sich bewegt, also in einen Schnittpunkt Q_r^* innerhalb \mathfrak{C}_r übergeht, so verlaufen die bei einer Operation (I) zunächst auftretenden Kegelschnitte^{*)} \mathfrak{R}^* bzw. ihre Q_r^* enthaltenden Äste \mathfrak{H}^* in der Umgebung des Punktes Q_r^* in $a(\mathfrak{H})$. Ist insbesondere $r = 1$, so liegt \mathfrak{R}^* bzw. \mathfrak{H}^* in der Umgebung von Q_1^* in $a(\mathfrak{H})$.

Somit sind, weil die vier festgehaltenen Punkte sämtlich Schnittpunkte und die einzigen gemeinsamen Punkte von \mathfrak{A} und \mathfrak{R}^* bzw. \mathfrak{A}^* sind, die Voraussetzungen von Nr. 2. 5. erfüllt. Es ist also \mathfrak{R}^* eine Hyperbel mit nicht kleinerer numerischer Exzentrizität als diejenige Hyperbel \mathfrak{H} , welche bei Beginn der Operation vorlag und um deren Ast \mathfrak{A} es sich handelte. Ist hingegen $r = 2t + 1$ mit $t > 0$, so liegt der von Q_{r-1} und von Q_r^* begrenzte, keine Q_r im Innern enthaltende Teilbogen von \mathfrak{R}^* in $a(\mathfrak{A})$, also der von Q_1 und Q_2 begrenzte Teilbogen \mathfrak{A}_1^* von \mathfrak{R}^* in $i(\mathfrak{A})$, also wieder eine zu \mathfrak{A}_1^* fremde Umgebung von Q_1 auf \mathfrak{R}^* in $a(\mathfrak{A})$, weil Q_1 (fester) Schnittpunkt von \mathfrak{A} und \mathfrak{R}^* . — Im Falle $r \equiv 0 \pmod{2}$ liegt \mathfrak{C}_r in $i(\mathfrak{A})$. Da wiederum Q_r bei einer Operation (I) ins Innere von \mathfrak{C}_r rückt, so liegt jetzt Q_r^* in $i(\mathfrak{A})$, also auch \mathfrak{A}_1^* in $i(\mathfrak{A})$, woraus die Beh. folgt.

2.4.1.2. Betr. die Operationen (II). Da es sich jeweils um eine beliebige kleine Änderung der vorliegenden Hyperbel handelt, so können alle diese Änderungen derart eingerichtet werden, daß die (durch alle zusammen bewirkte) Gesamtänderung der Exzentrizität beliebig klein wird, etwa kleiner als ein vorgegebener Bruchteil von $(e' - e)$, wobei e' die Exzentrizität der bei Beginn der Kontraktionsfolge zugrunde gelegten Hyperbel ist ($e < e'$). Soweit es von den Operationen (II) abhängt, bleiben dann die Exzentrizitäten aller auftretenden Kegelschnitte größer als e .

2.4.1.3. Betr. die Operationen (III). Bei diesen Operationen wird der jeweils vorliegende Kegelschnitt nicht geändert, also auch nicht die Exzentrizität.

2.5. Der in Nr. 2.4.1.1. benutzte Hilfssatz lautet:

Vor. Es sei \mathfrak{A} der eine Ast einer Hyperbel \mathfrak{H} . Ferner seien Q_1, \dots, Q_4 vier verschiedene (eigentliche) Punkte von \mathfrak{A} in natürlicher Anordnung¹⁴⁾ auf \mathfrak{A} . Weiter sei \mathfrak{R} ein Kegelschnitt derart, daß die Q_1, \dots, Q_4 auf einem beschränkten (zusammenhängenden) Teilbogen \mathfrak{T} von \mathfrak{R} liegen¹⁵⁾. Schließlich seien Q_2 und Q_3 enthalten in dem von Q_1 und Q_4 begrenzten offenen Teilbogen \mathfrak{T}_{14} von \mathfrak{T} ; und es liege eine zu \mathfrak{T}_{14} fremde Umgebung von Q_1 auf \mathfrak{R} im Äußern von \mathfrak{A} . — Beh. Es ist \mathfrak{R} eine Hyperbel, deren Exzentrizität nicht kleiner ist als die von \mathfrak{H} .

Wegen des (naheliegenden) Beweises kann auf eine an anderer Stelle¹⁶⁾ angegebene Verallgemeinerung des vorstehenden Hilfssatzes verwiesen werden.

¹⁴⁾ D. h. bei geeigneter Orientierung von \mathfrak{A} sollen die Q_1, \dots, Q_4 in dieser Anordnung auf \mathfrak{A} aufeinanderfolgen.

¹⁵⁾ Also auf dem gleichen Aste von \mathfrak{R} , falls \mathfrak{R} eine Hyperbel sein sollte.

¹⁶⁾ Vgl. Haupt, Bemerkungen über Konvexbogen. (Erscheint an anderer Stelle.)

Zum isoperimetrischen Problem für die nichteuklidischen Geometrien.

Von

Alexander Dinghas in Berlin.

1. *Einleitung.* In einer früheren Arbeit¹⁾, welche vor einiger Zeit in der *Math. Zeitschrift* erschienen ist, behandelte ich für den euklidischen, sphärischen und hyperbolischen Raum eine dem klassischen isoperimetrischen Problem äquivalente Aufgabe von Erhard Schmidt.

Unter Benutzung einer veränderlichen Krümmung und Verwendung Weierstraßscher homogener Koordinaten gelang es mir, diese Schmidtsche Aufgabe²⁾ für die drei obengenannten Räume folgendermaßen einheitlich zu formulieren:

Es bezeichnen x_1, x_2, \dots, x_{n+1} ($n \geq 2$) die Weierstraßschen homogenen Koordinaten des Raumes R_n mit

$$(1.1) \quad x_1^2 + x_2^2 + \dots + x_n^2 + \frac{x_{n+1}^2}{K} = \frac{1}{K} \quad (K \geq 0)$$

und es sei \mathfrak{H}_a der durch die Gleichung

$$(1.2) \quad x_1^2 + x_2^2 + \dots + x_{n-1}^2 \leq \frac{\sin^2 \sqrt{K} a}{K} \quad (a > 0)$$

dargestellte „Zylinder“. Es soll unter allen Körpern, die in (1.2) eingeschrieben werden können³⁾ und die eine vorgegebene Oberfläche besitzen, derjenige festgestellt werden, welcher das größte Volumen besitzt.

Diese Aufgabe wurde wie bei Erhard Schmidt⁴⁾ unter Beschränkung auf Rotationskörper gelöst, jedoch gelten, wie dies dort näher gezeigt

¹⁾ Zum isoperimetrischen Problem in Räumen konstanter Krümmung. *Math. Zeitschr.* 47 (1942), S. 677–737.

²⁾ I. c. 734.

³⁾ Dabei wird verstanden, daß die Projektion des Körpers parallel zur „Achse“ von (1.2) auf die Ebene $x_n = 0$ die Kugel $x_1^2 + x_2^2 + \dots + x_{n-1}^2 \leq \frac{\sin^2 \sqrt{K} a}{K}$ vollständig ausfüllt.

⁴⁾ Über das isoperimetrische Problem im Raum von n Dimensionen. *Math. Zeitschr.* 44 (1939), S. 689–788. Über die isoperimetrische Aufgabe im n -dimensionalen Raum konstanter negativer Krümmung. I. Die isoperimetrischen Ungleichungen in der hyperbolischen Ebene und für Rotationskörper in n -dimensionalen hyperbolischen Raum. Ebenda 46 (1940), S. 204–230. Die isoperimetrischen Ungleichungen auf der gewöhnlichen Kugel und für Rotationskörper in n -dimensionalen Raum. Ebenda 46 (1940),

wurde, die dort entwickelten Rechnungen für jeden in (1. 2) einbeschriebenen Körper, der auf dem Rand von (1. 2) einen ringsherum gehenden Randteil besitzt.

Wie am Schluß meiner genannten Arbeit gezeigt wurde, kann diese Aufgabe unabhängig vom Vorzeichen der Krümmung einheitlich behandelt werden und es ergibt sich, daß die Koeffizienten der erhaltenen Ungleichungen außer a nur noch von K und einem anderen wesentlichen Parameter abhängen, so daß der Übergang von einem dieser Räume in einen anderen durch stetige Variation der Krümmung erreicht werden kann.

Der Zweck der vorliegenden Arbeit besteht darin, ein den Schmidtschen Aufgaben ähnliches Problem unter Zugrundelegung der Cayleyschen Maßbestimmung zu behandeln, und zwar einheitlich für die elliptische, euklidische und hyperbolische Geometrie. Durch eine kleine Ergänzung kann man dann auch die sphärische Geometrie einordnen. Da nun dieses Problem sich mit demjenigen von Erhard Schmidt und damit auch mit dem klassischen Problem als äquivalent erweist, so ist dieses letzte Problem für alle vier Geometrien gelöst.

Diese neue Aufgabe, die wir ebenfalls nach Erhard Schmidt benennen wollen, formulieren wir folgendermaßen:

Es sei C_n ein Cayley-Kleinscher Raum mit der Metrik

$$(1.3) \quad ds^2 = \frac{(1 + K \sum_1^n u_i^2) \sum_1^n du_i^2 - K (\sum_1^n u_i du_i)^2}{(1 + K \sum_1^n u_i^2)^2}$$

gegeben, und es sei

$$(1.4) \quad \frac{u_1^2 + u_2^2 + \dots + u_{n-1}^2}{\frac{\operatorname{tg}^2 \sqrt{K} a}{K}} - K u_n^2 \leq 1$$

die Gleichung eines „Zylinders“, den wir mit S_a bezeichnen. Es soll unter allen Körpern der vorgegebenen Oberfläche O , die in (1. 4) einbeschrieben werden können, derjenige bestimmt werden, welcher das Maximum des Volumens V aufweist.

Diese Aufgabe wird hier zuerst für die Ebene und dann für den Raum gelöst. Die geometrische Form von (1. 4) ist je nach dem Vorzeichen von K ein einschaliges Rotations-Hyperboloid ($K > 0$), ein gewöhnlicher Zylinder

S. 740—794. Dazu kommt noch die vor kurzem erschienene, groß angelegte Abhandlung: Über eine neue Methode zur Behandlung einer Klasse isoperimetrischer Aufgaben im großen. Ebenda 47 (1942), S. 489—642. Diese letzte Arbeit nimmt insofern eine Sonderstellung ein, als es hier Herrn Erhard Schmidt gelang, eine einheitliche Darstellung seiner isoperimetrischen Ungleichungen für die Ebene nichteuklidischer Geometrien (mit Ausnahme der elliptischen) zu erhalten. Die Arbeiten von Erhard Schmidt werden im folgenden der Reihe nach mit I, II, III, IV zitiert.

($K = 0$), oder ein Rotations-Ellipsoid ($K < 0$). Die Methode, die hier befolgt wird, ist im Grunde identisch mit derjenigen Methode, die ich in meiner vorhin erwähnten Arbeit benutzt habe. Vom methodischen Standpunkt aus bietet also die vorliegende Arbeit außer dem hier behandelten elliptischen Fall und mancher Verschärfungen einiger isoperimetrischen Ungleichungen prinzipiell nicht viel Neues.

Jedoch wird der Leser viele Einzelheiten finden, die von Interesse sind. So habe ich den Weg über den Gaußschen Divergenzsatz, den ich in meiner früheren Arbeit eingeschlagen hatte, verlassen, obwohl er auf eine größere Mannigfaltigkeit von Zulassungskörpern anwendbar ist. Die hier angewandte Methode führt nur für Rotationskörper zum Ziel, und zwar über die Formeln (3. 7) und (3. 35), welche die Oberfläche bzw. das Volumen eines Rotationskörpers in einem n -dimensionalen Raum konstanter Krümmung geben. Die Zurückführung des allgemeinen isoperimetrischen Problems auf diesen speziellen Fall erfolgt, wie dies von Erhard Schmidt in einer bald erscheinenden Arbeit gezeigt werden wird, durch eine Erweiterung des Schwarzschen Symmetrisierungsverfahrens auf Riemannsche Räume konstanter Krümmung.

Was die weiteren Entwicklungen dieser Arbeit anbetrifft, so werden wir sehen, daß durch die Einführung der Weierstraßschen Koordination die hier zuletzt angeführte Aufgabe auf die erste zurückgeführt werden kann. Eine einfache Umrechnung der gefundenen Extremalen gibt dann unter Berücksichtigung der Bewegungsgruppen des Cayleyschen Raumes eine vollständige Lösung der isoperimetrischen Aufgabe für die nichteuklidischen Geometrien.

Zum Schluß möchte ich noch folgende Bemerkung hinzufügen: Während der ganzen Arbeit spielen die reellen Funktionen $\frac{\sin \sqrt{K} x}{\sqrt{K}}$, $\cos \sqrt{K} x$, $\frac{\operatorname{tg} \sqrt{K} x}{\sqrt{K}}$ ($K \geq 0$) eine wichtige Rolle und stellen für $K > 0$ gewöhnliche trigonometrische Funktionen und für $K < 0$ die entsprechenden hyperbolischen Funktionen $\frac{\operatorname{sh} \sqrt{-K} x}{\sqrt{-K}}$, $\operatorname{ch} \sqrt{-K} x$, $\frac{\operatorname{tgh} \sqrt{-K} x}{\sqrt{-K}}$ dar. Für $K = 0$ werden sie durch Stetigkeit definiert und werden entsprechend gleich x , 1 , x . Man könnte hier diese Funktionen als Funktionen, die von zwei Parametern abhängen, einführen und eventuell neue Bezeichnungen verwenden, um den Übergang über die imaginäre Einheit zu vermeiden. Ich habe jedoch vorgezogen, die obige Schreibweise zu behalten, mit der Verabredung, daß diese Funktionen bei negativem K durch die entsprechende hyperbolische ersetzt werden sollen und für $K = 0$ durch die Funktionen x bzw. 1 , bzw. x .

Die vorliegende Behandlung des isoperimetrischen Problems darf als elementar bezeichnet werden, und es zeigt sich, daß sich alles ohne die allgemeine Theorie der Variationsrechnung erledigen läßt. Daß hier im Grunde das Restglied der isoperimetrischen Ungleichung mit dem Weierstraßschen

E-Faktor identisch ist, wird dem Kenner einleuchten. Dagegen scheint die vorliegende Methode schwierig auf allgemeinere Probleme, die Erhard Schmidt in der vorhin mit IV bezeichneten Abhandlung behandelt hat, anwendbar zu sein. Lediglich das isoperimetrische Problem für die von ihm als Ringkörper bezeichneten Körper kann unter Abänderung der Aufgabe mit den hier eingeschlagenen Methoden angegriffen werden. Dies wird Gegenstand einer späteren Arbeit bilden.

Kapitel I.

Die isoperimetrische Aufgabe für die zwei-dimensionalen nichteuklidischen Geometrien.

2. *Die Schmidtsche Aufgabe für eine Klasse zwei-dimensionaler symmetrischer Mannigfaltigkeiten.* Wir betrachten eine Mannigfaltigkeit \mathfrak{M} mit der Metrik

$$(2.1) \quad ds^2 = du^2 + \varphi^2(u)dv^2,$$

und setzen voraus, daß $\varphi(u)$ im Intervall

$$(2.2) \quad -A < u < A$$

stetig differenzierbare Funktion von u ist, und daß in jedem abgeschlossenen Teilintervall von (2.2)

$$(2.3) \quad \varphi(u) > 0$$

gilt. Das Variationsintervall von v lassen wir zunächst unbestimmt. Dies ändert sich je nach der Voraussetzung über die Mannigfaltigkeit \mathfrak{M} und kann, wie z. B. auf der hyperbolischen Ebene, auch unendlich sein. Von unserer Funktion $\varphi(u)$, von deren Eigenschaften die Struktur von \mathfrak{M} abhängt, fordern wir, daß sie folgende Bedingungen erfüllt:

1. Symmetriebedingung. Es gilt in (2.2)

$$(2.4) \quad \varphi(-u) = \varphi(u).$$

2. Wachstumsbedingung. Bezeichnet $\psi(u)$ das Integral von $\varphi(u)$

$$(2.5) \quad \psi(u) = \int_0^u \varphi(x) dx,$$

so ist die Funktion

$$(2.6) \quad \omega(u) = \frac{\psi(u)}{\varphi(u)}$$

eine in (2.2) monotone wachsende Funktion von u . Insbesondere ist in jedem Punkt dieses Intervalls $\omega' > 0$ ⁵⁾.

⁵⁾ Diese Bedingung kann durch die schwächere $\omega'(-a) > 0$, $\omega'(a) > 0$ ersetzt werden. Dabei ist a ($a > 0$) eine gleich zu definierende Zahl.

Die Mannigfaltigkeiten, die wir zugrunde legen, weisen also, wenn wir von der Wachstumsbedingung absehen, die Symmetrieeigenschaft auf. Zu diesen gehören z. B. alle Mannigfaltigkeiten konstanter Krümmung.

Für diese symmetrischen Mannigfaltigkeiten formulieren wir zuerst die Schmidtsche Aufgabe⁶⁾ folgendermaßen:

Unter allen einfach geschlossenen Kurven C auf \mathfrak{M} , welche im Streifen \mathfrak{M}_a (2.7)

$$-a \leq u \leq a \quad (-A < -a < a < A)$$

einbeschrieben werden können und die eine vorgegebene Länge L besitzen, soll diejenige bestimmt werden, die das Maximum des eingeschlossenen Inhalts F aufweist.

Als einbeschrieben bezeichnen wir jede einfach geschlossene Kurve, die in (2.7) verläuft und die mit den Parameterlinien $u = \pm a$ gemeinsame Punkte hat.

Um unsere Aufgabe zu lösen, führen wir auf \mathfrak{M}_a zwei Kurvenscharen F_1, F_2 mit den Gleichungen $v^* \pm F(u, \gamma) = c$ ein, wobei c den Scharparameter bedeutet und γ einen (jedesmal festen) zweiten Parameter darstellt. Wie man sich leicht überzeugt, geht die eine Familie aus der anderen hervor durch Spiegelung an der u -Achse, d. h. an der Parameterlinie $v = 0$. Wegen der gemachten Voraussetzung kann man noch verlangen, daß die Kurven $v^* \pm F(u, \gamma) = 0$ durch die beiden Punkte $u = \pm a, v = 0$ gehen und daß die Kurve $v^* + F(u, \gamma) = 0$ oberhalb der u -Achse verläuft. Wir bezeichnen die aus diesen beiden Kurven gebildete geschlossene Figur mit K_γ . Ferner folgt aus der gewählten Metrik unmittelbar, daß die Gesamtheit der Kurven F_1 bzw. F_2 kongruent sind, und daß sie durch die Transformation $v^* = v^* + C$ in sich übergehen.

Bevor wir nun die Familien F_1 bzw. F_2 explizit als „Extremalen“ angeben, bestimmen wir eine „Winkelkorrespondenz“ zwischen einer zulässigen [nämlich in (2.5) einbeschriebenen] Kurve C und den Kurven F_1 bzw. F_2 , und zwar so: Die Kurve C möge durch zwei Punkte $P_1(-a, q_1), P_2(a, q_2)$ in zwei Teile zerlegt werden, den „oberen“ Teil C_1 (der den Punkt mit der maximalen Ordinate v enthält) und den „unteren“ Teil C_2 . Ist nun $P(u, v)$ ein Punkt von C_1 , so ordnen wir diesem den Punkt $\bar{P}(u, v^*)$ ($v^* \geq 0$) von $v^* + F(u, \gamma) = 0$ zu. Wir definieren als Winkel (n, \bar{n}) den Winkel zwischen den äußeren Normalen n von C in P und \bar{n} von K_γ in \bar{P} . Da P, \bar{P} im allgemeinen nicht zusammenfallen, so wird als Winkel (n, \bar{n}) der Winkel definiert, den die äußere Normale n von C in C mit der Normalen derjenigen Kurve von F_1 bildet, welche durch P geht. Entsprechend wird der Winkel (n, \bar{n}) definiert, wenn P dem Teil C_2 angehört. An Stelle der Kurvenschar F_1 tritt dann die-

⁶⁾ Vgl. die unter IV zitierte Arbeit. Es handelt sich hier um einen sehr speziellen Fall der Schmidtschen Aufgabe, der aber zur Lösung der klassischen isoperimetrischen Aufgabe vollkommen ausreicht.

jenige von F_2 ein. Es handelt sich also bei der Zuordnung $P \rightarrow \bar{P}$ um eine Korrespondenz, die beim ersten Blick etwas Willkürliches besitzt, es kann aber leicht gezeigt werden, daß dieses willkürliche Element (die Punkte P_1, P_2) bei der Integration verschwindet.

Einige Worte noch über die äußeren Normalen. Die äußere Normale \mathfrak{n} von C wird so definiert, daß beim Durchlaufen von C in positivem Sinne (entgegengesetzt dem Uhrzeiger auf der Ebene!) der Bereich nach rechts gerichtet ist. Für die Kurven F_1 bzw. F_2 wird diese Richtung entsprechend definiert und zwar so, daß diese Verabredung erhalten bleibt, wenn man eine Kurve von F_1 (oberhalb $v = 0$) und eine Kurve von F_2 (unterhalb $v = 0$) mit den Parameterlinien zu einer geschlossenen Kurve ergänzt. Wir fordern noch, daß, wenn man zwei Kurven derselben Familie zum Übereinstimmen bringt, die Orientierung beider Kurven dieselbe sei.

Nach diesen Vorbereitungen legen wir die Figur K_γ zugrunde und betrachten das Integral⁷⁾

$$(2.10) \quad J = \int \cos(n, \bar{n}) ds$$

längs einer zulässigen Kurve C . Dabei ist ds das Bogenelement von C . Drücken wir die u, v eines Punktes P von C als Funktion der Bogenlänge s von C aus⁸⁾, so sind die Koordinaten von \mathfrak{n}

$$(2.11) \quad \cos(n, u) = \varphi \frac{du}{ds}, \quad \cos(n, v) = -\frac{dv}{ds}.$$

Wir definieren jetzt unsere Kurvenscharen F_1, F_2 durch die Differentialgleichungen

$$(2.12) \quad \begin{aligned} \cos(\bar{n}, u) &= \varphi \frac{dv^*}{ds} = k_1 \omega(u), \\ \cos(\bar{n}, v) &= -\frac{du}{ds} = \pm \sqrt{1 - k_1^2 \omega^2(u)}, \end{aligned}$$

wobei (u, v^*) die laufenden Koordinaten eines Punktes P von F_1 bzw. F_2 bedeuten. Das $+$ -Zeichen entspricht der Kurvenschar F_1 , während das

⁷⁾ Solche Integrale betrachtete auch Darboux (Théorie générale des Surfaces III, Paris 1894, 1. Aufl., S. 140) und Schwarz (vgl. Gesammelte Mathem. Abhandlungen, Bd. 1, 1890, S. 332. Berlin, Springer). Als ich meine vorhin erwähnte Arbeit schrieb, kannte ich weder diese Stelle bei Darboux (ich habe von dort nur die Gleichung der geodätischen Kreise entnommen), noch die zitierte Stelle bei Schwarz. Auf diesen letzteren bin ich erst auf Anregung von Herrn Prof. Erhard Schmidt gekommen. Freilich beschäftigen sich weder Darboux noch Schwarz mit der Schmidtschen Problemstellung, sie erhalten auch aus diesem Grunde die hier entwickelten isoperimetrischen Gleichungen nicht. Darboux hat neben einigen verwandten isoperimetrischen Aufgaben als Hauptaufgabe, eine Formel für die geodätische Krümmung zu beweisen, die auf Bonnet zurückgeht.

⁸⁾ Wir nehmen durchweg in der ganzen Arbeit an, daß $u(s), v(s)$ totalstetige Funktionen von s sind. Zugrunde wird dann immer das Lebesguesche Integral gelegt.

—Zeichen für die Kurven der Schar F_2 vorbehalten bleiben soll. Der Parameter k_γ wird hier gleich $\frac{1}{\varrho_\gamma} = k_0 \cos \gamma$ (mit $0 \leq 2\gamma < \pi$) gesetzt und

$$(2.13) \quad k_0 = \frac{1}{\varrho_0} = \frac{1}{\omega(a)}.$$

Durch Division der beiden Formeln (2.12) erhält man als Differentialgleichung der Kurvenschar F_1 ,

$$(2.14) \quad \varphi \frac{dv^*}{ds} + \frac{k_\gamma \omega}{\sqrt{1 - k_\gamma^2 \omega^2}} = 0$$

und als Differentialgleichung der Kurvenschar F_2

$$(2.15) \quad \varphi \frac{dv^*}{ds} - \frac{k_\gamma \omega}{\sqrt{1 - k_\gamma^2 \omega^2}} = 0.$$

Die Kurven von F_1 bzw. F_2 , welche durch die beiden Punkte $u = \pm a, v = 0$ gehen, bezeichnen wir mit K_1, K_2 und das durch sie begrenzte Gebiet mit \mathfrak{R}_γ . Zur Definition des von einer Kurve C_γ begrenzten Gebietes \mathfrak{R}_γ sei hier nur dieses gesagt, daß dies so definiert ist, daß beim Durchlaufen von C_γ in dem vorhin definierten positiven Sinne \mathfrak{R}_γ links bleiben soll.

Wie man aus (2.14), (2.15) unmittelbar entnimmt, lauten die Gleichungen von K_1, K_2

$$(2.16) \quad v^* \pm \int_{-a}^u \frac{k_\gamma \omega dx}{\varphi \sqrt{1 - k_\gamma^2 \omega^2}} = 0$$

und das Pluszeichen entspricht der Kurve K_1 .

Die Bedeutung von γ ergibt sich aus der ersten Formel (2.12) unter Berücksichtigung der Fortsetzung (2.13).

Es folgt dann aus diesen beiden Formeln

$$\cos(n, u) = \frac{\omega(u)}{\omega(a)} \cos \gamma,$$

woraus für $u = +a$

$$(2.17) \quad \cos(n, u) = \cos \gamma$$

folgt. γ bedeutet also den Winkel, den die Normalen von K_1 bzw. K_2 mit der u -Achse ($v = 0$) in dem Punkt $u = a$ bilden. Es braucht hier kaum wiederholt zu werden, daß die Kurven der beiden Familien F_1, F_2 durch die Gleichungen

$$(2.18) \quad \begin{aligned} v^* + \int_{-a}^u \frac{k_\gamma \omega dx}{\varphi \sqrt{1 - k_\gamma^2 \omega^2}} &= c, \\ v^* - \int_{-a}^u \frac{k_\gamma \omega dx}{\varphi \sqrt{1 - k_\gamma^2 \omega^2}} &= c' \end{aligned}$$

gegeben werden, wobei c, c' willkürliche Konstante bedeuten. Die beiden Familien gehen offenbar durch Spiegelung ineinander über.

Nun kehren wir zu (7.10) zurück, und entwickeln unter Benutzung von (2.11) und (2.12). Es wird dann wegen

$$(2.19) \quad \cos(n, \bar{n}) = \cos(n, u) \cos(\bar{n}, v) + \cos(n, v) \cos(\bar{n}, u),$$

$$(2.20) \quad \cos(n, \bar{n}) = k_\gamma \varphi \omega \frac{dv}{ds} \mp \sqrt{1 - k_\gamma^2 \omega^2} \frac{du}{ds},$$

und zwar ist hier das Minusvorzeichen zu nehmen, wenn der Punkt u, v auf C_1 läuft, und das Plusvorzeichen, wenn dieser Punkt auf C_2 läuft. Daraus folgt durch Integration über die ganze Kurve C

$$(2.21) \quad \int_C \cos(n, \bar{n}) ds = k_\gamma \int_C \varphi \omega \frac{dv}{ds} ds + \int_C \mp \sqrt{1 - k_\gamma^2 \omega^2} \frac{du}{ds} ds.$$

Das erste Integral bedeutet hier, wie man leicht einsieht, den Inhalt F des von C begrenzten Bereiches. Das zweite Integral ist von der speziellen Kurve unabhängig und gleich

$$(2.22) \quad 4 \int_0^a \sqrt{1 - k_\gamma^2 \omega^2} du.$$

Wir erhalten dann aus (2.21)

$$(2.23) \quad \int_C \cos(n, \bar{n}) ds = k_\gamma F + 4 \int_0^a \sqrt{1 - k_\gamma^2 \omega^2} du.$$

Schreibt man $1 - 2 \cos^2 \frac{(n, \bar{n})}{2}$ an Stelle von $\cos(n, \bar{n})$, so folgt daraus

$$(2.24) \quad L - k_\gamma F = 4 \int_0^a \sqrt{1 - k_\gamma^2 \omega^2} du + 2 R(C, C_\gamma)$$

mit

$$(2.25) \quad R(C, C_\gamma) = \int_C \sin^2 \frac{(n, \bar{n})}{2} ds \geq 0.$$

Daß man aus dieser Formel alle Ergebnisse über das isoperimetrische Problem für die hier betrachteten Mannigfaltigkeiten erhalten kann, habe ich schon in meiner vorhin zitierten Arbeit gezeigt. Dort wurde auch gezeigt, daß die Wahl der Teilungspunkte P_1, P_2 von C auf das Endresultat nicht von Einfluß ist.

Wir geben hier die Ableitung der beiden fundamentalen isoperimetrischen Gleichungen wieder: Fall I. $0 < \gamma < \frac{\pi}{2}$. Die aus den zwei Bögen $K_1 K_2$ bestehende Extremale C_γ heißt eine Extremale vom linsenförmigen Typus. Bezeichnet L_γ die Länge von C_γ und F_γ das von ihr begrenzte Gebiet, so

folgt, wenn man die früheren Rechnungen mit K_γ an Stelle von C wiederholt, wegen $R(C_\gamma, C_\gamma) = 0$

$$(2.26) \quad L_\gamma - k_\gamma F_\gamma = 4 \int_0^a \sqrt{1 - k_\gamma^2 \omega^2} du$$

und daraus durch Kombination mit (2.24)

$$(2.27) \quad L - L_\gamma = k_\gamma (F - F_\gamma) + 2 R(C, C_\gamma),$$

oder

$$(2.28) \quad F - F_\gamma = k_\gamma (L - L_\gamma) - 2 e_\gamma R(C, C_\gamma).$$

Fall II. $\gamma = 0$. In diesem Falle ist $e_0 = \omega(a)$ und wird

$$(2.29) \quad F - F_0 = \omega(a) (L - L_0) - 2 \omega(a) R(C, C_0)$$

mit

$$(2.30) \quad F_0 = 4 \int_0^a \frac{\omega^3 dx}{\sqrt{\omega^2(a) - \omega^2(x)}}, \quad L_0 = 4 \omega(a) \int_0^a \frac{dx}{\sqrt{\omega^2(a) - \omega^2(x)}}.$$

Nun ziehen wir die beiden symmetrischen Hälften, aus denen C_0 besteht, in zwei Teile eines Teleskops auseinander und bezeichnen die so entstandene Kurve, welche aus den beiden symmetrischen Hälften C_1, C_2 von C_0 und den zwei Bögen der Parallelkreise $u = \pm a$, mit C_q^* besteht. Dabei bedeutet q die Differenz der v -Koordinaten der Berührungspunkte von C_1, C_2 mit einem der Parallelkreise $u = \pm a$ z. B. mit $u = a$. Eine Extremale C_q^* nennen wir eine Extremale vom Kreis-Rechteck-Typus. Ihr Inhalt wird mit F_q^* und ihre Länge mit L_q^* bezeichnet. Da ihrer Konstruktion gemäß die Gleichung

$$(2.31) \quad F_0 - \omega(a) L_0 = F_q^* - \omega(a) L_q^*$$

besteht, so können wir an Stelle von (2.29) die Gleichung

$$(2.32) \quad F - F_q^* = \omega(a) (L - L_q^*) - 2 \omega(a) R(C, C_q^*)$$

setzen, da das Restglied, genommen in bezug auf die neue Extremale C_q^* , gleich dem alten bleibt. Aus diesen Gleichungen folgt die Lösung der Schmidtschen Aufgabe unmittelbar, wenn man zu jedem $L > 4a$ die entsprechende Extremale mit $L_\gamma = L$ bzw. $L_q^* = L$ konstruiert. Daß es nur eine Extremale gibt, ist für die linsenförmigen Extremalen aus der leicht beweisbaren Formel⁹⁾

$$(2.33) \quad L_\gamma = 4 \int_0^a \frac{du}{\sqrt{1 - k_\gamma^2 \omega^2}}$$

⁹⁾ Vgl. meine vorhin zitierte Arbeit, S. 702.

zu entnehmen, während die Monotonie von L_q^* als Funktion von q trivial ist und aus der Darstellung

$$(2.34) \quad L_q^* = L_0 + 2q\varphi(a)$$

sofort abzulesen ist.

Da er für das Folgende von Wichtigkeit ist, bringen wir den allgemeinen Satz wieder und verweisen für weitere Einzelheiten auf die vorhin zitierte Arbeit, in der die Rechnungen ausführlicher dargestellt werden. Die Antwort auf die vorhin gestellte Aufgabe von Erhard Schmidt lautet:

Die Lösung der Schmidtschen Aufgabe für die zugrunde gelegten symmetrischen Mannigfaltigkeiten ist für

$$(2.35) \quad 4a < L < L_0$$

eine linsenförmige Extremale von der Länge $L_v = L$. Gilt ferner

$$(2.36) \quad \eta_0 = \int_0^a \frac{\omega \, dx}{\varphi \sqrt{\omega^2(a) - \omega^2(x)}} < \pi,$$

so sind die Lösungen der Aufgabe für

$$(3.37) \quad L_0 \leq L < L_0 + 2(2\pi - 2\eta_0)\varphi(a)$$

Extremalen vom Kreis-Rechteck-Typus mit $L_q^ = L$.*

Die Voraussetzung (2.36) wird verlangt, wenn man Extremalen fordert, die ein schlichtes Gebiet begrenzen. Für die Spezialfälle $\varphi(u) = \cos \sqrt{K}u$ ($K > 0$) ist diese Forderung identisch erfüllt und für die beiden anderen Fälle ($K \leq 0$) ist sie nicht notwendig, da v dort von $-\infty$ bis $+\infty$ variiert.

Für $L \leq 4a$ ist die Aufgabe unlösbar. Somit haben wir das wichtigste entwickelt, was wir für die späteren Anwendungen unbedingt brauchen.

Kapitel II.

Anwendung auf die nichteuklidischen Geometrien.

3. *Vorbereitende Tatsachen. Das Cayleysche Schema der nichteuklidischen Geometrie.* Wir schicken zuerst die wichtigsten Tatsachen voraus über dasjenige Schema, das Cayley¹⁰⁾ der nichteuklidischen Geometrie gegeben hat.

¹⁰⁾ Die Literatur zur nichteuklidischen Geometrie ist sehr groß. Hier kann nur das wiederholt werden, was zum Verständnis der nachfolgenden Ausführungen unbedingt notwendig ist. Für frühere Entwicklungen vgl. man außer der Kleinschen Abhandlung „Über die sogenannte nichteuklidische Geometrie“ [Math. Annalen 4 (1871), S. 573–625], das Buch von Cartan „Leçons sur la géométrie des espaces de Riemann“. Paris, Gauthier-Villars, 1928, und die Note von Appel am Schluß des 5. Bandes seiner Mécanique rationelle. Paris, Gauthier-Villars, 1925.

Für den ebenen Fall, der uns hier beschäftigt, bedient sich dieses Schema als „Unterlage“ der gewöhnlichen euklidischen Ebene π , auf der sich eine Kurve zweiten Grades Q , die „Fundamentalkurve“, befindet. Je nach der Natur von π unterscheidet man zwei Fälle:

1. Q ist rein imaginär mit reeller Gleichung (Pseudoreell). Dann gehören alle Punkte von π dem Schema an und die Geometrie umfaßt alle Punkte dieser Ebene.

2. Q ist eine reelle konvexe Kurve zweiten Grades.

Die Geometrie umfaßt dann nur das Innere von Q . Die gewöhnliche euklidische Geometrie kann dann bekanntlich als Grenzfall einer dieser Geometrien betrachtet werden.

Wie man leicht beweisen kann, ist es immer möglich, Q auf die Form eines imaginären bzw. reellen Kreises zurückzuführen. Aus diesem Grunde können wir von vornherein annehmen, daß die Fundamentalkurve in bezug auf ein Achsenkreuz XOY die Gleichung

$$(3.1) \quad X^2 + Y^2 + \frac{1}{K} = 0$$

hat, mit $K \geq 0$. $K > 0$ entspricht dann der elliptischen Geometrie, $K < 0$ der hyperbolischen und der Grenzfall $K \rightarrow 0$ (ausgeartete Q) der gewöhnlichen euklidischen Geometrie.

Sind nun $P(X, Y)$, $P'(X + dX, Y + dY)$ zwei benachbarte Punkte unseres Schemas, so wird die Entfernung ds dieser beiden Punkte durch die Formel definiert

$$(3.2) \quad ds^2 = \frac{(1 + K(X^2 + Y^2))(dX^2 + dY^2) - K(X dX + Y dY)^2}{(1 + K(X^2 + Y^2))^2},$$

wobei diese Metrik im Falle der elliptischen Geometrie überall mit Ausnahme von unendlich großen X, Y gilt.

Somit sind die wichtigsten Grundtatsachen, die zum Verständnis des Späteren erforderlich sind, wiederholt. Einige topologische Schwierigkeiten bei der Formulierung der Aufgabe im elliptischen Fall werden an der richtigen Stelle genauer präzisiert werden.

Wir führen jetzt folgende Koordinatentransformation ein:

$$(3.3) \quad \begin{aligned} X &= \frac{1}{\sqrt{K}} \frac{\operatorname{tg} \sqrt{K} u}{\cos \sqrt{K} v}, \\ Y &= \frac{1}{\sqrt{K}} \operatorname{tg} \sqrt{K} v, \end{aligned}$$

wobei für $K > 0$, u, v zuerst im Intervall

$$(3.4) \quad -\frac{\pi}{2\sqrt{K}} \leq u \leq \frac{\pi}{2\sqrt{K}}$$

variieren und für $K \leq 0$ im Intervall $-\infty$ bis $+\infty$.

Eine elementare Rechnung liefert dann für das Bogenelement ds , ausgedrückt in den neuen Koordinaten,

$$(3.5) \quad ds^2 = du^2 + \cos^2 \sqrt{K} u dv^2.$$

Hiermit ist unsere Maßbestimmung mittels (3.3) auf die vorhin untersuchte Form

$$(3.6) \quad ds^2 = du^2 + q^2(u) dv^2$$

zurückgeführt worden.

Ist nun $K < 0$, so wird durch (3.3) das Innere von (3.1) eindeutig auf die uv -Ebene abgebildet. Für $K \rightarrow 0$ artet diese Transformation in eine Identität aus und sie braucht uns nicht näher zu beschäftigen. Es bleibt also nur der Fall übrig, dem wir uns jetzt zuwenden werden.

Zuerst wird der endliche Teil von π durch (3.3) auf das Quadrat

$$(3.7) \quad -\frac{\pi}{2\sqrt{K}} < u, v < \frac{\pi}{2\sqrt{K}}$$

abgebildet werden, und zwar eindeutig. Wir ergänzen dieses Quadrat durch Hinzufügung von zwei Rändern zum Quadrat G_0

$$(3.8) \quad \begin{aligned} -\frac{\pi}{2\sqrt{K}} &\leq u < \frac{\pi}{2\sqrt{K}}, \\ -\frac{\pi}{2\sqrt{K}} &\leq v < \frac{\pi}{2\sqrt{K}}, \end{aligned}$$

und bezeichnen dieses als Fundamentalbereich. Durch die Abbildung (3.3) geht dann das durch das Hyperbelpaar

$$(3.9) \quad X^2 - Y^2 \operatorname{tg}^2 \sqrt{K} a = \frac{\operatorname{tg}^2 \sqrt{K} a}{K}$$

begrenzte Gebiet G^* in das Rechteck G_a von G_0

$$(3.10) \quad -a \leq u \leq a, \quad -\frac{\pi}{2\sqrt{K}} \leq v < \frac{\pi}{2\sqrt{K}}$$

über. Da wir uns hier hauptsächlich für diejenigen Bewegungen der XY -Ebene interessieren, bei welchen G^* in sich übergeht, so konstruieren wir die universelle Überlagerungsfläche des Rechtecks G_a oder des Bandes

$$(3.11) \quad -a \leq u < a, \quad -\frac{\pi}{2\sqrt{K}} \leq v < \frac{\pi}{2\sqrt{K}}.$$

Dazu spiegeln wir zuerst den Fundamentalbereich G_0 an den Seiten AB , $B'A'$. Durch Wiederholung dieses Verfahrens an den entsprechenden horizontalen Seiten der so entstandenen Nebengebiete erhält man einen Streifen, der aus unendlich vielen kongruenten Quadraten $\dots G_{-n}, G_{-n+1}, \dots$,

$G_{-1}, G_0, G_1 \dots G_n \dots$ (Fig. 1) besteht. Analytisch können diese durch die Ungleichungen

$$(3.12) \quad -\frac{\pi}{2\sqrt{K}} \leq u < \frac{\pi}{2\sqrt{K}}, \quad -\left(n + \frac{1}{2}\right) \frac{\pi}{2\sqrt{K}} \leq v < \left(n + \frac{1}{2}\right) \frac{\pi}{2\sqrt{K}}$$

dargestellt werden, wenn dabei n eine ganze nicht negative Zahl bedeutet.

Die Ränder dieser Bereiche sind wie beim bekannten Möbiusschen Band zugeordnet. So sind z. B. die Punkte von AB den Punkten von $B'A'$ zugeordnet, nach vorheriger Spiegelung der Strecke AB an der v -Achse, wie dies in der Figur gezeichnet wird. Obwohl wir später auf die elliptischen Bewegungen nochmals zurückkommen werden, versuchen wir, diese Tatsache mit Hilfe der projektiven Koordinaten zu veranschaulichen. Dabei benutzen wir folgende Tatsachen, die wir erst in Nr. 11 beweisen werden: Die elliptische Ebene kann durch drei Projektive oder Weierstraßsche Koordinaten x_1, x_2, x_3 dargestellt werden. Diese sind durch die Gleichung

$$(3.13) \quad x_1^2 + x_2^2 + \frac{x_3^2}{K} = \frac{1}{K} \quad (K > 0)$$

verknüpft. Setzt man dann

$$(3.14) \quad X = \frac{x_1}{x_3}, \quad Y = \frac{x_2}{x_3},$$

so lautet die entsprechende Metrik

$$(3.15) \quad ds^2 = dx_1^2 + dx_2^2 + \frac{dx_3^2}{K}.$$

Durch die Substitution $\sqrt{K} x_3' = x_3$ kann man die Größe K aus (3.13) (3.15) teilweise eliminieren. Man erhält dann an Stelle dieser Gleichungen

$$(3.16) \quad x_1^2 + x_2^2 + x_3'^2 = \frac{1}{K},$$

$$(3.16') \quad ds^2 = dx_1^2 + dx_2^2 + dx_3'^2.$$

Den elliptischen Raum kann man¹¹⁾ dadurch erhalten, daß man die Punkte

¹¹⁾ Alle diese Tatsachen sind bekannt und werden nur wegen einer größeren Einheitlichkeit der Darstellung wiederholt. Zu dem nachfolgenden Schema vgl. am besten Klein, Vorlesungen über nichteuklidische Geometrien, S. 151. Berlin, Springer, 1928. Bei der Gelegenheit möchte ich noch bemerken, daß die Aufgabe für die elliptische Geometrie, so wie diese hier gestellt ist, im Grunde identisch ist mit der isoperimetrischen Aufgabe auf einer offenen Halbkugel vom Radius $\frac{1}{\sqrt{K}}$.

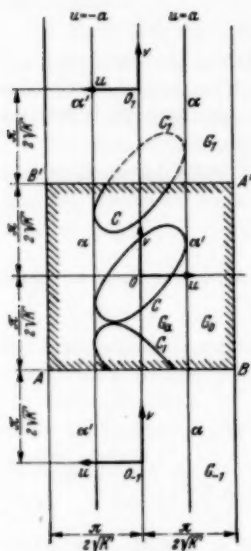


Fig. 1.

der Kugel (3. 16) von ihrem Mittelpunkt aus auf die Punkte derjenigen Ebene E projiziert, die (3. 16) im Punkte $x'_3 = +1$ berührt. Denn wählt man auf E die X, Y -Achsen parallel zu x_1 bzw. x_2 , so erhalten die Projektionspunkte $P(\pm x_1, \pm x_2, \pm x_3)$ die Koordinaten X, Y .

Die Transformation (3. 3) erhält man dann, wenn man auf der Kugel Polarkoordinaten einführt, und zwar durch

$$(3. 17) \quad \begin{aligned} x_1 &= \frac{1}{\sqrt{K}} \sin \sqrt{K} u, \\ x_2 &= \frac{1}{\sqrt{K}} \cos \sqrt{K} u \sin \sqrt{K} v, \\ x'_3 &= \frac{1}{\sqrt{K}} \cos \sqrt{K} u \cos \sqrt{K} v. \end{aligned}$$

Unser Fundamentalgebiet ist nichts anderes als das Netz der u, v -Linien der oberen Halbkugel von (3. 16).

Zwei diametral-gegenüberliegende Punkte auf (3. 16) entsprechen nur einem Punkte auf E . Man kann diese Zweideutigkeit aufheben, indem man die untere Halbkugel von (3. 16) wegläßt und dann zwei diametral-gegenüberliegende Punkte des Äquators $x_3 = 0$ als einen einzigen Punkt¹²⁾ betrachtet.

Definitionsgemäß sind Bewegungen der elliptischen Ebene Drehungen der Kugel (3. 16). Bei uns kommen dann nur solche Bewegungen in Frage, die x_1 festlassen, also nur Rotationen in der $x_2 x'_3$ -Ebene. Diesen entsprechen die Schiebungen in der uv -Ebene längs der v -Achse. Um die Verknüpfung der Bereiche $\dots G_{-2}, G_{-1}, G_0, G_1, G_2 \dots$ einzusehen, braucht man nur die Bewegung des starren, die Kugel (3. 16) berührenden Achsenkreuzes zu verfolgen, das sich auf dieser stetig bewegt und das so liegt, daß die eine Achse desselben in der $x_2 x'_3$ -Ebene liegt.

Wir begnügen uns mit diesen Andeutungen. Ähnliche Tatsachen gelten auch in höherdimensionalen elliptischen Räumen, worauf wir noch teilweise später zurückkommen werden.

Wir schließen hier nur diese Ausführungen mit dem Hinweis, daß der Zusammenhang von G^* bzw. G_a mit dem Zusammenhang eines Möbiusschen Bandes identisch ist. Diese Tatsache ist einleuchtend und kann mittels des sphärischen Schemas unmittelbar nachgewiesen werden. G^* ist ja nichts anderes als das Gebiet der Oberfläche von (3. 16), das zwischen den Parallelkreisen $x_1 = \pm \frac{\sin \sqrt{K} a}{\sqrt{K}}$ liegt.

¹²⁾ Klein, l. c. [Fußnote 11)], S. 152. Hiermit wird auch klar, daß die Strecke BB' als Bild des Punktes $x_1 = \pm \frac{1}{K}$ zu betrachten ist.

Mit Hilfe der Transformation (3. 3) geht unsere in der Einleitung gestellte Aufgabe für $n = 2$ in ein äquivalentes Problem in der uv -Ebene über. Wegen der Definition des von einer geschlossenen Kurve begrenzten Bereiches soll hier der Wortlaut dieser Aufgabe nochmals genau wiederholt werden.

Ist C eine einfache Kurve auf der XY -Ebene, deren Punkte alle im Endlichen liegen, so definiert sie zwei Bereiche, von denen der eine (dessen Punkte alle im Endlichen liegen) einfach zusammenhängend ist, während der andere (der „komplementäre“) zweifach zusammenhängend ist. Unter einem von einer einfachen geschlossenen Kurve begrenzten Bereich werden wir immer denjenigen verstehen, dessen Zusammenhang einfach ist. Als Kurven C werden alle Kurven zugelassen werden, die durch eine nichteuklidische Bewegung in eine solche zurückgeführt werden, deren Punkte im Endlichen liegen.

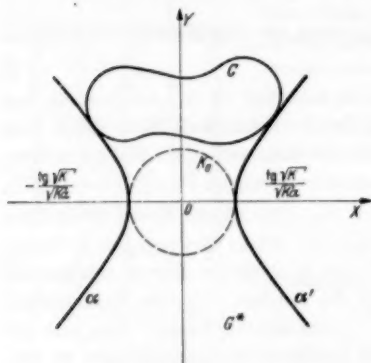


Fig 2.

Wir formulieren nun unsere Aufgabe für die elliptische Ebene folgendermaßen:

Es soll in der XY -Ebene unter allen im „Streifen“ G^*

$$(3. 18) \quad X^2 - \operatorname{tg}^2 \sqrt{K} a Y^2 \leq \frac{\operatorname{tg}^2 \sqrt{K} a}{K}$$

einbeschriebenen geschlossenen einfachen Kurven C von der elliptischen Länge L diejenige bestimmt werden, welche ein Bereich mit dem maximalen elliptischen Inhalt F begrenzt.

Wie wir schon vorhin erwähnt haben, brauchen alle Punkte einer

zugelassenen Kurve nicht im Endlichen zu liegen. Es wird aber verlangt, daß es möglich wird, C durch eine „Parallelverschiebung“ längs der Y -Achse in eine „kongruente“ Kurve C' mit dieser Eigenschaft zurückzuführen (Fig. 2). Ferner werden L vorläufig keine Einschränkungen auferlegt. Vielmehr werden sich die Bedingungen dazu durch die geforderte Schlichtheit der Extremalgebiete zwangsläufig ergeben.

Die so gestellte Aufgabe ist nun völlig äquivalent mit folgender Aufgabe in der uv -Ebene:

Unter allen einfach geschlossenen Kurven C , die im Streifen G_0

$$(3. 19) \quad -a \leq u \leq a, \quad -\frac{\pi}{2\sqrt{K}} \leq v < \frac{\pi}{2\sqrt{K}}$$

des Fundamentalbereiches G_0 einbeschrieben sind, soll diejenige bestimmt werden, welche bei vorgegebener Länge L das Maximum des eingeschlossenen elliptischen Inhalts F aufweist.

Man kann allerdings ohne weiteres die zweite Forderung (3. 11) fallen lassen und dieselbe Aufgabe für den Überlagerungsstreifen lösen. Das ist gleichbedeutend mit dem Verzicht auf eine schlichte Überdeckung der elliptischen Ebene durch den von C begrenzten Bereich.

4. *Der sphärische und hyperbolische Fall.* Nach dieser topologisch-metrischen Vorbereitung für den elliptischen Fall kehren wir zu unserer Aufgabe zurück und formulieren sie für den sphärischen und hyperbolischen Fall.

Liegt zuerst im dreidimensionalen euklidischen Raum die Kugel \mathfrak{S} vor mit der Gleichung (in bezug auf ein Achsenkreuz $Oxyz$)

$$(4.1) \quad x^2 + y^2 + z^2 = \frac{1}{K} \quad (K > 0),$$

so erfüllen wir diese Gleichung durch

$$(4.2) \quad \begin{aligned} x &= \frac{1}{\sqrt{K}} \sin \sqrt{K} u, \\ y &= \frac{1}{\sqrt{K}} \cos \sqrt{K} u \sin \sqrt{K} v, \\ z &= \frac{1}{\sqrt{K}} \cos \sqrt{K} u \cos \sqrt{K} v, \end{aligned}$$

wobei u, v in den Intervallen

$$(4.3) \quad -\frac{\pi}{2\sqrt{K}} \leq u \leq \frac{\pi}{2\sqrt{K}}, \quad -\frac{\pi}{\sqrt{K}} \leq v < \frac{\pi}{\sqrt{K}}$$

variieren. Die universelle Überlagerungsfläche der in den beiden Punkten $(\pm 1, 0, 0)$ punktierten Kugel \mathfrak{S} kann dann durch den in der Fig. 3 angegebenen Streifen dargestellt werden. Der Übergang zu den Nebenbereichen G_1, G_2, \dots bzw. G_{-1}, G_{-2}, \dots geschieht im Gegensatz zu dem elliptischen Fall ohne die Spiegelung an der v -Achse. Ein vollständiger Umlauf eines Punktes P auf der Kugel längs eines durch ihn gehenden Parallelkreises $x = \text{const.}$ entspricht auf der uv -Ebene einer Verschiebung längs der v -Achse um $\frac{2\pi}{\sqrt{K}}$.

Analytisch definieren wir den Fundamentalbereich G_0 von \mathfrak{S} durch die Ungleichungen

$$(4.4) \quad -\frac{\pi}{2\sqrt{K}} < u < \frac{\pi}{2\sqrt{K}}, \quad -\frac{\pi}{\sqrt{K}} \leq v < \frac{\pi}{\sqrt{K}}$$

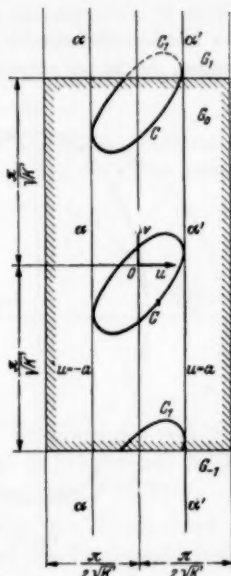


Fig. 3.

und nehmen als Metrik (die man aus $ds^2 = dx^2 + dy^2 + dz^2$ umrechnet)

$$(4.5) \quad ds^2 = du^2 + \cos^2 \sqrt{K} u dv^2.$$

Für den sphärischen Fall formulieren wir dann die Aufgabe folgendermaßen¹³⁾:

Unter allen einfach geschlossenen Kurven, welche im Streifen

$$(4.6) \quad -a \leq u \leq a, \quad -\frac{\pi}{\sqrt{K}} \leq v < \frac{\pi}{\sqrt{K}}$$

des Fundamentalbereiches G_0 (4.3) eingeschrieben sind, soll diejenige bestimmt werden, welche bei vorgegebener sphärischer Länge L das Maximum des eingeschlossenen sphärischen Inhalts F erreicht.

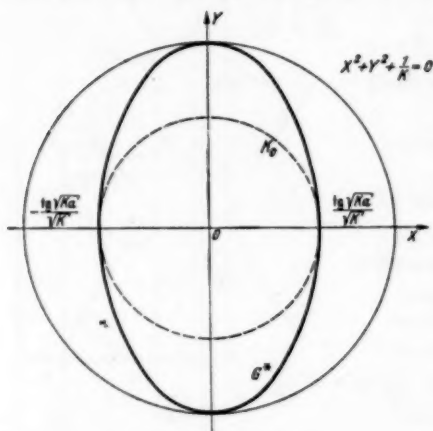


Fig. 4.

Ebenso wie im elliptischen Fall kann die zweite Forderung (4.6) weggelassen werden. Die schlichte Überdeckung braucht dann nicht mehr stattzufinden.

Zum Schluß fügen wir einige Worte zum hyperbolischen Fall hinzu:

Die Schmidtsche Aufgabe erhält hier folgenden Wortlaut: *Unter allen einfach geschlossenen Kurven, welche in dem aus den beiden Halbellipsen*

$$(4.7) \quad X^2 - \operatorname{tg}^2 \sqrt{K} Y^2 = \frac{\operatorname{tg}^2 \sqrt{K} a}{K} \quad (K < 0)$$

begrenzten „Streifen“ G^* eingeschrieben sind, soll diejenige bestimmt werden, welche unter vorgegebener hyperbolischer Länge L das Maximum des eingeschlossenen hyperbolischen Inhalts F aufweist. Wie hier ausdrücklich hervorgehoben werden möchte, handelt es sich bei (4.7) (Fig. 4) nicht um eine Ellipse, sondern um zwei symmetrische Ellipsenhälften, welche den Fundamentalkreis in den Punkten $(0, \pm \frac{1}{\sqrt{-K}})$ berühren. Dies folgt mehr durch projektive Überlegung, wird jedoch uns hier nicht näher beschäftigen.

¹³⁾ Vgl. auch Dinghaas, l. c. S. 685.

Nun geht durch die Transformation (3. 3) das Innere des Fundamentalkreises in die endliche uv -Ebene über und das Innere von G^* in den Streifen

$$(4. 8) \quad -a \leq u \leq a.$$

Somit erhält unsere Aufgabe in der uv -Ebene folgende Formulierung:

Unter allen einfach geschlossenen Kurven C von vorgegebener Länge L , die im Streifen (4. 7) der hyperbolischen uv -Ebene mit der Metrik $ds^2 = du^2 + \cos^2 \sqrt{K} u dv^2$ ($K < 0$) eingeschrieben sind, soll diejenige ermittelt werden, welche das Maximum des eingeschlossenen Inhalts F aufweist.

Nach diesen allgemeinen Vorbereitungen werden in der nächsten Nummer alle diese Aufgaben einheitlich behandelt werden.

5. Eine allgemeine isoperimetrische Gleichung. Durch die Ausführungen der vorigen Nummer ist die in der Einleitung für den Fall der Ebene gestellte Aufgabe mit der nachfolgenden äquivalent:

Es sei in der uv -Ebene mit der Metrik

$$(5. 1) \quad ds^2 = du^2 + \cos^2 \sqrt{K} u dv^2$$

im Streifen G_a^{14})

$$(5. 2) \quad -a \leq u \leq a$$

gegeben¹⁵⁾. Es soll unter allen zulässigen in G_a eingeschriebenen einfach geschlossenen Kurven C von vorgegebener Länge L ($L > 2a$) diejenige bestimmt werden, welche das Maximum des eingeschlossenen Inhalts F aufweist.

Ich habe in der vorhin erwähnten Abhandlung eine Lösung dieser von Erhard Schmidt stammenden Aufgabe gegeben, und zwar für den euklidischen, sphärischen und hyperbolischen Fall, der auf anderen Grundlagen, als die von Erhard Schmidt gegebenen Beweise, fußt. Wenn ich nun die Lösung dieses Problems hier unter allgemeineren Gesichtspunkten wiederhole, so geschieht dies, weil man durch die vorherigen Entwicklungen auch den elliptischen Fall mit einbeziehen kann. Dabei erfährt die frühere Darstellung gewisse Verkürzungen und eine größere Vereinheitlichung.

Zur Lösung dieser Aufgabe kehren wir nun zu den Ausführungen von Nr. 2 zurück und stellen fest, daß die dort gestellten Forderungen für die

¹⁴⁾ Es wird sich zuerst für $K > 0$ um den Streifen des Fundamentalbereiches G_0 handeln.

¹⁵⁾ Für $K > 0$ soll $a < \frac{\pi}{2\sqrt{K}}$ sein.

Funktion $\varphi(u)$ hier erfüllt sind. Die Differentialgleichung der Extremalen F_1, F_2 , die durch $(\pm a, 0)$ gehen, lautet:

$$(5.4) \quad \cos \sqrt{K} u \frac{dv}{du} \pm \frac{\frac{\operatorname{tg} \sqrt{K} u}{\sqrt{K}}}{\sqrt{\varrho_\gamma^2 - \frac{\operatorname{tg}^2 \sqrt{K} u}{K}}} = 0.$$

Dabei ist ϱ_γ der geodätische Krümmungsradius des zugehörigen Bogens des geodätischen Kreises.

Setzt man ferner

$$(5.5) \quad \varrho_\gamma = \frac{\varrho_0}{\cos \gamma} = \frac{\operatorname{tg} \sqrt{K} a}{\sqrt{K} \cos \gamma},$$

so ist γ der Winkel, unter dem F_1 (bzw. F_2) die u -Achse in dem Punkte $u = a$ schneidet.

Wir behaupten jetzt:

Für $K > 0$ lautet die Gleichung der durch (5.4) definierten Extremalen F_1, F_2

$$(5.6) \quad \sin^2 \sqrt{K} u + \cos^2 \sqrt{K} u \sin^2 \sqrt{K} (v \pm \eta_0) = \frac{K \varrho_\gamma^2}{1 + K \varrho_\gamma^2},$$

wobei

$$(5.7) \quad \frac{\sin \sqrt{K} \eta_0}{\sqrt{K}} = \frac{\varrho_\gamma \sin \gamma}{\sqrt{1 + K \varrho_\gamma^2}}$$

ist. Der Fall $K = 0$ ist hier in (5.6) als Grenzfall enthalten und man findet aus (5.6) (5.7) als Gleichung der Extremalen

$$(5.8) \quad u^2 + (v \pm a \operatorname{tg} \gamma)^2 = \frac{a^2}{\cos^2 \gamma},$$

d. h. die zwei bekannten Kreisbögen vom Radius $\frac{a}{\cos \gamma}$.

Für $K < 0$ sind drei Fälle zu unterscheiden, je nachdem $1 + K \varrho_\gamma^2$ größer, gleich oder kleiner als Null ist. Ist $1 + K \varrho_\gamma^2 > 0$, so lautet die Gleichung der Extremalen

$$(5.9) \quad \sin^2 \sqrt{K} u + \cos^2 \sqrt{K} u \sin^2 \sqrt{K} (v \pm \eta_0) = \frac{K \varrho_\gamma^2}{1 + K \varrho_\gamma^2},$$

wobei η_0 wieder durch (5.7) bestimmt wird. Für $1 + K \varrho_\gamma^2 < 0$ werden die Extremalen durch

$$(5.10) \quad \sin^2 \sqrt{K} u + \cos^2 \sqrt{K} u \cos^2 \sqrt{K} (v \pm \eta_0) = \frac{K \varrho_\gamma^2}{1 + K \varrho_\gamma^2}$$

gegeben. Dabei wird noch η_0 durch

$$(5.11) \quad \cos \sqrt{K} \eta_0 = \sqrt{\frac{K \varrho_\gamma^2}{1 + K \varrho_\gamma^2}} \sin \gamma$$

bestimmt. Die Gleichung der Extremalen für $1 + K e_\gamma^2 = 0$, welche durch Grenzübergang aus (5. 9) oder (5. 10) erhalten werden kann, lautet

$$(5.12) \quad \cos \sqrt{K} u \left(\cos \sqrt{K} v \pm \frac{\sqrt{-K} \sin \sqrt{K} v}{\sqrt{K}} \right) = \cos \sqrt{K} a.$$

Die Gleichung (5. 6) stellt, wie man leicht sehen kann, geodätische Kreise dar, die durch die Punkte $(\pm a, 0)$ gehen. Die Gleichung (5. 9) stellt eine Schar von hyperbolischen Kreisen dar, die durch dieselben Punkte gehen, während (5. 10) die Gesamtheit aller Abstandskurven (Hyperzyklen) darstellt, welche ebenfalls durch $(\pm a, 0)$ gehen.

Für die Darstellung der Extremalen für $K < 0$ kann man ebenso gut hyperbolische Funktionen benutzen, obwohl wir dies hier nicht systematisch betreiben werden.

So findet man an Stelle von (5. 9), (5. 10) und (5. 12) auch die Darstellung ($K' = -K > 0$)

$$(5.13) \quad \operatorname{sh}^2 \sqrt{K'} u + \operatorname{ch}^2 \sqrt{K'} u \operatorname{sh}^2 \sqrt{K'} (v \pm \eta_0) = \frac{K' e_\gamma^2}{1 - K' e_\gamma^2},$$

$$(5.14) \quad -\operatorname{sh}^2 \sqrt{K'} u + \operatorname{ch}^2 \sqrt{K'} u \operatorname{ch}^2 \sqrt{K'} (v \pm \eta_0) = \frac{-K'}{1 - K' e_\gamma^2}$$

und

$$(5.15) \quad \operatorname{ch} \sqrt{K'} u (\operatorname{ch} \sqrt{K'} v \pm \operatorname{sh} \sqrt{K'} v) = \operatorname{ch} \sqrt{K'} a.$$

Jedoch wird darauf nicht näher eingegangen werden. Um nun nachzuweisen, daß (5. 6), (5. 9) und (5. 12) Integrale von (5. 4) darstellen, verfahren wir am einfachsten so, wobei wir uns zuerst der Einfachheit halber auf den oberen Bogen von (5. 6) beschränken:

Wegen

$$(5.16) \quad \begin{aligned} \sin \sqrt{K} u &= \frac{\operatorname{tg} \sqrt{K} u}{\sqrt{1 + \operatorname{tg}^2 \sqrt{K} u}}, \\ \cos \sqrt{K} u &= \frac{1}{\sqrt{1 + \operatorname{tg}^2 \sqrt{K} u}} \end{aligned}$$

folgt durch Einsetzen in die (5. 6)

$$(5.17) \quad \frac{\sin \sqrt{K} (v + \eta_0)}{\sqrt{K}} = \sqrt{e_\gamma^2 - \frac{\operatorname{tg}^2 \sqrt{K} u}{K}} \cdot \sqrt{\frac{1}{1 + K e_\gamma^2}}.$$

Daraus folgt

$$(5.18) \quad \cos \sqrt{K} (v + \eta_0) \frac{dv}{du} = - \frac{1}{\sqrt{1 + K e_\gamma^2}} \frac{\operatorname{tg} \sqrt{K} u}{\sqrt{K} \cos^2 \sqrt{K} u} \frac{1}{\sqrt{e_\gamma^2 - \frac{\operatorname{tg}^2 \sqrt{K} u}{K}}}.$$

Da nun eben wegen (5. 17)

$$(5. 19) \quad \cos \sqrt{K} (v + \eta_0) = \sqrt{1 - \sin^2 \sqrt{K} (v + \eta_0)} = \frac{1}{\sqrt{1 + K e_\gamma^2 \cos \sqrt{K} u}}$$

ist, so folgt durch Einsetzen

$$(5. 20) \quad \cos \sqrt{K} u \frac{dv}{du} + \frac{\frac{\operatorname{tg} \sqrt{K} u}{\sqrt{K}}}{\sqrt{e_\gamma^2 - \frac{\operatorname{tg}^2 \sqrt{K} u}{K}}} = 0,$$

also die Differentialgleichung einer Kurve von der Schar F_1 . Man kann auch rückwärts von (5. 20) ausgehend, unter Voraussetzung von $1 + K e_\gamma^2 > 0$, zu der Gleichung (5. 6) gelangen. Ebenso leicht bestätigt man, daß die Gleichung

$$(5. 21) \quad \sin^2 \sqrt{K} u + \cos^2 \sqrt{K} u \sin^2 \sqrt{K} (v - \eta_0) = \frac{K e_\gamma^2}{1 + K e_\gamma^2}$$

eine durch $(\pm a, 0)$ gehende Extremale ist, welche die Differentialgleichung

$$(5. 22) \quad \cos \sqrt{K} u \frac{dv}{du} - \frac{\frac{\operatorname{tg} \sqrt{K} u}{\sqrt{K}}}{\sqrt{e_\gamma^2 - \frac{\operatorname{tg}^2 \sqrt{K} u}{K}}} = 0$$

erfüllt. Denn diese geht durch die Transformation $v' = -v$ in die Gleichung von F_1 über.

Für $K < 0$ muß man nun auch den Fall $1 + K e_\gamma^2 \leq 0$ berücksichtigen. Ist zuerst $1 + K e_\gamma^2 < 0$, so folgt aus (5. 9) für die Extremale, welche oberhalb $v = 0$ verläuft,

$$(5. 23) \quad \cos \sqrt{K} (v + \eta_0) = \sqrt{\frac{K}{1 + K e_\gamma^2}} \sqrt{e_\gamma^2 - \frac{\operatorname{tg}^2 \sqrt{K} u}{K}}.$$

Daraus ergibt sich durch Differentiation

$$(5. 24) \quad \sqrt{K} \sin \sqrt{K} (v + \eta_0) \frac{dv}{du} = - \sqrt{\frac{K}{1 + K e_\gamma^2}} \frac{\frac{\operatorname{tg} \sqrt{K} u}{\sqrt{K}}}{\cos^2 \sqrt{K} u} \frac{1}{\sqrt{e_\gamma^2 - \frac{\operatorname{tg}^2 \sqrt{K} u}{K}}}.$$

Da nun aus (5. 23) die Gleichung

$$(5. 25) \quad \sin \sqrt{K} (v + \eta_0) = \frac{\sqrt{1 + \operatorname{tg}^2 \sqrt{K} u}}{\sqrt{1 + K e_\gamma^2}} = \frac{1}{\sqrt{1 + K e_\gamma^2 \cos \sqrt{K} u}}$$

folgt, so wird mit Rücksicht auf (5. 24)

$$(5. 26) \quad \cos \sqrt{K} u \frac{dv}{du} = - \frac{\frac{\operatorname{tg} \sqrt{K} u}{\sqrt{K}}}{\sqrt{e_\gamma^2 - \frac{\operatorname{tg}^2 \sqrt{K} u}{K}}},$$

d. h. wieder die Differentialgleichung der Extremalen. Ganz analog verfährt man für den unterhalb $v = 0$ verlaufenden Teil von (5. 10). Es ändert sich nur das Vorzeichen der rechten Seite von (5. 21).

Für $1 + K \varrho_\gamma^2 \rightarrow 0$ wird die Gleichung der Extremalen:

$$(5. 27) \quad \frac{dv}{du} \pm \sqrt{-K} \frac{\operatorname{tg} \sqrt{K} u}{\sqrt{K}} = 0,$$

und man kann leicht verifizieren, daß (5. 12) dieser Differentialgleichung genügt. (5. 12) wird dann die Lösung sein, welche durch $(\pm a, 0)$ geht.

Zum Schluß wäre auch vielleicht nochmals nötig zu betonen, daß in (5. 6), (5. 8), (5. 9), (5. 10) und (5. 12) sowie in der ganzen Arbeit nur diejenigen Teile der angegebenen Gleichungen als Extremalbögen zu betrachten sind, welche durch $(\pm a, 0)$ gehen, zur u -Achse spiegelbildlich liegen und ganz im gegebenen Streifen liegen.

6. *Ein isoperimetrischer Satz.* Wir bezeichnen nun wie in Nr. 2 mit C_γ den Rand des von den Extremalen (5. 6) oder (5. 10) oder (5. 12) begrenzten Gebietes R_γ , mit L_γ die Länge von C_γ und mit F_γ den Inhalt von C_γ , jedesmal berechnet in der jeweiligen Metrik.

Man erhält dann für alle vier nichteuklidischen Geometrien die fundamentale Gleichung

$$(6. 1) \quad F - F_\gamma = \varrho_\gamma (L - L_\gamma) - 2 \varrho_\gamma R(C, C_\gamma).$$

Extremalen von dem Typus (5. 6), (5. 10) oder (5. 12) habe ich in der vorhin zitierten Abhandlung Extremalen von linsenförmigem Typus genannt. Sie entsprechen dem Typus, den Erhard Schmidt in einer kürzlich erschienenen großen Abhandlung¹⁶⁾ Extremalen vom Typus (A) genannt hat. Die Größe $R(C, C_\gamma)$ stellt ein Restglied dar, das durch (2. 17) gegeben wird und das nur dann verschwindet, wenn C mit C_γ gestaltlich übereinstimmt.

Für $\gamma = 0$ ist $\eta_0 = 0$ und die Extremalen haben die Gleichung

$$(6. 2) \quad \sin^2 \sqrt{K} u + \cos^2 \sqrt{K} u \sin^2 \sqrt{K} v = \frac{K \varrho_0^2}{1 + K \varrho_0^2} = \sin^2 \sqrt{K} a.$$

Sie stellt einen geodätischen Kreis K_0 dar, der den Streifen G_a in den Punkten $(\pm a, 0)$ berührt.

Wir schneiden nun K_0 in den Punkten $(\pm a, 0)$ durch und ziehen die so entstehenden Halbkreisbögen K_1, K_2 wie zwei Teile eines Teleskops auseinander, etwa in einer Entfernung q . Es entsteht dann (wenn die neue Lage von K_1, K_2 symmetrisch zu der v -Achse angenommen wird), ein von K_1, K_2 und den zwei Strecken

$$(6. 3) \quad u = \pm a, \quad -\frac{q}{2} \leq v \leq \frac{q}{2}$$

¹⁶⁾ Schmidt V., S. 501.

begrenzter Bereich B_q^* , dessen Rand wir mit C_q^* bezeichnen. Da nun

$$(6.4) \quad F_0 - \varrho_0 L_0 = F_q^* - \varrho_0 L_q^*$$

ist und noch

$$(6.5) \quad R(C, C_q^*) = R(C, C_0)$$

gilt, so erhalten wir aus (6.1) für $\gamma = 0$ die Gleichung

$$(6.6) \quad F - F_q^* = \varrho_0(L - L_q^*) - 2 \varrho_0 R(C, C_q^*).$$

F_q^* bedeutet hier den Inhalt von B_q^* und L_q^* die Länge von C_q^* . Extremalen vom Typus C_q^* habe ich in meiner erwähnten Abhandlung Extremalen vom Kreis-Rechteck-Typus genannt. Sie entsprechen den Extremalen, welche Erhard Schmidt Extremale vom Typus (B, 3) genannt hat.

Aus diesen beiden Gleichungen (6.1) und (6.6) folgt unmittelbar der Beweis folgenden Satzes¹⁷⁾, dessen Inhalt schon auf Erhard Schmidt¹⁸⁾ zurückgeht:

Die in 5. vorgelegte Aufgabe für den Fundamentalstreifen G_0 besitzt eine Lösung, wenn

$$(6.7) \quad L > 2a$$

ist. Ist dann

$$(6.8) \quad 2a < L < L_0,$$

so ist die Lösung eine Extremale C_γ von linsenförmigem Typus mit $L_\gamma = L$. Für

$$(6.9) \quad L_0 \leq L$$

und $K \leq 0$ ist die Extremale C_q^ vom Kreis-Rechteck-Typus mit $L_q^* = L$.*

Für $K > 0$ tritt wegen der Beschränktheit von L nach oben eine Einschränkung für q ein nach oben, und zwar gilt für das Maximum q_m von q

$$(6.10) \quad 2a + q_m < \frac{\pi}{\sqrt{K}}$$

für die elliptische und

$$(6.11) \quad 2a + q_m < \frac{2\pi}{\sqrt{K}}$$

für die sphärische Geometrie.

Setzt man dann

$$(6.12) \quad \bar{q} = 2\left(\frac{\pi}{2\sqrt{K}} - a\right) \quad \text{bzw.} \quad \bar{q} = 2\left(\frac{\pi}{\sqrt{K}} - a\right),$$

je nachdem die elliptische bzw. die sphärische Geometrie gilt, so gibt es für

$$(6.13) \quad L_0 \leq L < L_{\bar{q}}^*$$

¹⁷⁾ Vgl. Dinghas, l. c. S. 685.

¹⁸⁾ Für genaue Literatur vgl. man Dinghas, l. c. S. 688.

immer eine Extremale C_q^* vom Kreis-Rechteck-Typus, welche im Fundamentalbereich G^* liegt.

Man kann natürlich die Forderung der einfachen Bedeckung fallen lassen. Dann braucht C_q nicht notwendig im Fundamentalbereich G^* zu liegen und die Einschränkung (6. 13) fällt fort.

Was die explizite Auswertung von L_0 angeht, so findet man leicht

$$(6. 14) \quad L_0 = 2\pi \frac{\sin \sqrt{K} a}{\sqrt{K}}.$$

Entsprechend findet man für die Größtwerte L_q^*

$$(6. 15) \quad L_q^* = L_0 + 2 \cos \sqrt{K} a \left(\frac{\pi}{2\sqrt{K}} - a \right)$$

für die elliptische Geometrie, und

$$(6. 15') \quad L_q^* = L_0 + 2 \cos \sqrt{K} a \left(\frac{\pi}{\sqrt{K}} - 2a \right)$$

für die sphärische Geometrie.

Der Beweis dieses Satzes folgt unmittelbar aus (6. 1) und (6. 6), wenn man die Extremale C_γ^* bzw. C^* mit $L = L_\gamma$ bzw. $L = L_q^*$ konstruiert. Daß zu einem gegebenen $L = L_\gamma$ nur eine solche Extremale existiert, folgt leicht aus der Monotonie von L_γ als Funktion von γ . Berechnet man nämlich die Länge L von C unter Zugrundelegung von (5. 1) und der Differentialgleichung (5. 4), so findet man leicht

$$(6. 16) \quad L_\gamma = 4 \int_0^a \frac{du}{\sqrt{1 - k_\gamma^2 \frac{\operatorname{tg}^2 \sqrt{K} u}{K}}} \quad \left(k_\gamma = \frac{1}{e_\gamma} \right),$$

woraus die Monotonie von L_γ unmittelbar folgt. Die Monotonie von C_q^* ist allerdings trivial, da ja nach Definition

$$(6. 17) \quad L_q^* = L_0 + 2q \cos \sqrt{K} a$$

ist.

Hiermit ist der Satz bewiesen. Denn konstruiert man die durch L bestimmte Extremale mit $L = L_\gamma$ bzw. L_q^* , so folgt aus den vorhin bewiesenen Gleichungen

$$(6. 18) \quad F - F_\gamma \leq 0 \quad \text{bzw.} \quad F - F_q^* \leq 0$$

und das Gleichheitszeichen kann nur dann eintreten, wenn das zugehörige Restglied verschwindet. Das tritt aber nur ein, wenn C durch Parallelverschiebung aus C_γ bzw. C_q^* hervorgeht.

7. Die Gestalt der Extremalen auf der Cayleyschen Ebene. Wie sehen nun unsere Extremalen in der XY -Ebene aus? Nach den Formeln (3.3) ist

$$(7.1) \quad \begin{aligned} X &= \frac{1}{\sqrt{K}} \frac{\operatorname{tg} \sqrt{K} u}{\cos \sqrt{K} v}, \\ Y &= \frac{1}{\sqrt{K}} \operatorname{tg} \sqrt{K} v. \end{aligned}$$

Man könnte nun durch Substitution in den Formeln (5.9), (5.10) und (5.12) die Gleichungen der linsenförmigen Extremalen erhalten. Dieser Weg ist aber rechnerisch mühsam und wenig übersichtlich. Wir werden deswegen einen anderen Weg einschlagen, der über die homogenen Koordinaten zum Resultat führt und der zugleich die Struktur der Substitution (3.3) von anderer Seite her beleuchtet.

Wir setzen für die drei nichteuklidischen Geometrien¹⁹⁾

$$(7.2) \quad x_1^2 + x_2^2 + \frac{x_3^2}{K} = \frac{1}{K} \quad (K \geq 0)$$

und erfüllen diese Gleichungen durch

$$(7.3) \quad \begin{aligned} x_1 &= \frac{1}{\sqrt{K}} \sin \sqrt{K} u, \\ x_2 &= \frac{1}{\sqrt{K}} \cos \sqrt{K} u \sin \sqrt{K} v, \\ x_3 &= \cos \sqrt{K} u \cos \sqrt{K} v. \end{aligned}$$

Dann ist

$$(7.4) \quad X = \frac{x_1}{x_3}, \quad Y = \frac{x_2}{x_3}.$$

Da nun die u, v Koordinaten die Mittelpunkte des durch (5.3) dargestellten Kreises $u = 0, v = \mp \eta_0$ sind, so folgt mit

$$(7.5) \quad \bar{x}_1 = 0, \quad \bar{x}_2 = \frac{1}{\sqrt{K}} \sin \sqrt{K} \eta_0, \quad \bar{x}_3 = \cos \sqrt{K} \eta_0$$

$$(7.6) \quad x_1^2 + (x_2 \bar{x}_3 \mp x_3 \bar{x}_2)^2 = \frac{e_\gamma^2}{1 + K e_\gamma^2} \quad (1 + K e_\gamma^2 > 0).$$

Ebenfalls folgt aus (5.10)

$$(7.7) \quad K x_1^2 + (K x_2 \bar{x}_3 \mp x_3 \bar{x}_2)^2 = \frac{K e_\gamma^2}{1 + K e_\gamma^2} \quad (1 + K e_\gamma^2 < 0).$$

¹⁹⁾ Die sphärische Geometrie wird hier beiseite gelassen, da ja für diesen Fall die Gestalt der Extremalen unmittelbar zu haben ist. Es handelt sich hier um gewöhnliche Kreise auf der Kugel, welche durch die beiden Punkte $(\pm a, 0)$ gehen.

Diese Koordinaten, die für $K > 0$ schon in Nr. 3 eingeführt wurden, heißen auch homogene Koordinaten von Weierstraß. Vgl. darüber am besten das schon in der Fußnote ¹⁸⁾ zitierte Buch von Cartan.

Den Fall (5.12) behandeln wir zum Schluß als Grenzfall.

Erledigen wir den Fall $K \geq 0$, also (7.7). Dazu drücken wir x_1, x_2, x_3 durch X, Y aus. Da man bekanntlich $x_3 > 0$ wählen kann²⁰⁾, so finden wir

$$(7.8) \quad \begin{aligned} x_1 &= \frac{X}{\sqrt{1+K(X^2+Y^2)}}, \\ x_2 &= \frac{Y}{\sqrt{1+K(X^2+Y^2)}} \end{aligned}$$

und

$$(7.9) \quad x_3 = \frac{1}{\sqrt{1+K(X^2+Y^2)}}.$$

Es wird also mit

$$(7.10) \quad X_0 = \frac{\bar{x}_1}{\bar{x}_3}, \quad Y_0 = \frac{\bar{x}_2}{\bar{x}_3} > 0$$

$$(7.11) \quad X^2 + \frac{(Y \pm Y_0)^2}{1+K Y_0^2} = \frac{e_\gamma^2}{1+K e_\gamma^2} (1 + K(X^2 + Y^2)).$$

Wie man leicht nachprüft, stellen die beiden Kurvenfamilien (7.11) eine Schar von Kurven zweiter Ordnung dar, welche durch die Punkte

$$(7.12) \quad P_1\left(\frac{\operatorname{tg} \sqrt{K} a}{\sqrt{K}}, 0\right), \quad P_2\left(-\frac{\operatorname{tg} \sqrt{K} a}{\sqrt{K}}, 0\right)$$

gehen.

Für $K < 0$ finden wir ebenfalls (7.11), sofern $1 + K e_\gamma^2 > 0$ ist, mit dem Unterschied, daß hier

$$(7.13) \quad -\frac{1}{\sqrt{-K}} < Y_0 < \frac{1}{\sqrt{-K}}$$

ist. Für $1 + K e_\gamma^2 < 0$, also im Falle einer Abstandskurve, folgt aus (7.8)

$$(7.14) \quad K X^2 + \frac{(1 \pm K Y Y_0)^2}{1+K Y_0^2} = \frac{K e_\gamma^2}{1+K e_\gamma^2} (1 + K(X^2 + Y^2))$$

mit Y_0 in (7.13). Setzt man nun

$$(7.15) \quad K = -K', \quad K' Y_0 Y'_0 = 1,$$

so wird

$$(7.16) \quad X^2 + \frac{(Y \pm Y'_0)^2}{1+K' Y'_0{}^2} = \frac{e_\gamma^2}{1+K' e_\gamma^2} (1 + K(X^2 + Y^2))$$

mit

$$(7.17) \quad |Y'_0| > \frac{1}{\sqrt{-K'}}.$$

²⁰⁾ Cartan, l. c. S. 164.

Wegen $1 + K e_\gamma^2 < 0$ und (7.17) stellen (7.16) eine Familie von Ellipsen dar, welche durch (7.12) gehen. Wie steht es mit dem Fall $e_\gamma = \frac{1}{\sqrt{-K}}$? Um dies zu beantworten, transformieren wir (7.11) und (7.16) mit Hilfe von (5.7) und (5.11). Man erhält, indem wir Y_0 an Stelle von Y'_0 schreiben und die Einschränkung (7.13) für $K < 0$ fallen lassen, eine einzige Gleichung für die linsenförmigen Extremalen, nämlich folgende:

$$(7.18) \quad X^2(1 + K Y_0^2) + (Y \pm Y_0)^2 = \frac{e_\gamma^2}{1 + K e_\gamma^2} (1 + K(X^2 + Y^2)).$$

Die Gleichung (7.18) stellt allerdings alle linsenförmigen Extremalen dar, unabhängig vom Vorzeichen der Krümmung K . Wir können diese auch so transformieren, daß wir nämlich für $1 + K e_\gamma^2 > 0$ die Gleichung

$$(7.19) \quad \frac{\sin \sqrt{K} \eta_0}{\sqrt{K}} = \frac{e_\gamma \sin \gamma}{\sqrt{1 + K e_\gamma^2}} = \frac{1}{\sqrt{K}} \frac{\operatorname{tg} \sqrt{K} \eta_0}{\sqrt{1 + \operatorname{tg}^2 \sqrt{K} \eta_0}} = \frac{Y_0}{\sqrt{1 + K Y_0^2}}$$

nach Y_0 auflösen und den entsprechenden Wert in (7.18) einsetzen. Für $1 + K e_\gamma^2 < 0$ wird wegen (5.11)

$$(7.20) \quad \cos \sqrt{K} \eta_0 = \sqrt{\frac{K e_\gamma^2}{1 + K e_\gamma^2}} \sin \gamma = \frac{1}{\sqrt{1 + K Y_0^2}} = \frac{\sqrt{K Y_0^2}}{\sqrt{1 + K Y_0^2}}.$$

Schreiben wir dann auch im Falle (7.24) Y_0 an Stelle von Y'_0 , so ergibt sich, wie man leicht feststellt, in allen Fällen unabhängig vom Vorzeichen von K

$$(7.21) \quad Y_0 = \frac{\sin \sqrt{K} a}{\sqrt{K}} \operatorname{tg} \gamma.$$

Durch Einführung in die (7.18) erhalten wir dann die allgemeine Darstellung der linsenförmigen Extremalen

$$(7.22) \quad X^2(1 + \sin^2 \sqrt{K} a \operatorname{tg}^2 \gamma) + \left(Y \pm \frac{\sin \sqrt{K} a}{\sqrt{K}} \operatorname{tg} \gamma\right)^2 = \frac{\sin \sqrt{K} a}{K \cos^3 \gamma} (1 + K(X^2 + Y^2)).$$

Diese Gleichung kann man leicht auch in die Form

$$(7.23) \quad \frac{X^2}{A} + \frac{(Y - \bar{Y})^2}{B} = 1$$

bringen, wobei die Größen A, B und \bar{Y} durch die Formeln

$$(7.24) \quad A = \frac{1}{K} \frac{\sin^2 \sqrt{K} a}{\cos^2 \sqrt{K} a - \sin^2 \gamma},$$

$$B = \frac{1}{K} \frac{\sin^2 \sqrt{K} a \cos^2 \sqrt{K} a \cos^2 \gamma}{(\cos^2 \sqrt{K} a - \sin^2 \gamma)^2}$$

und

$$(7.25) \quad \bar{Y} = \frac{1}{\sqrt{K}} \frac{\sin \sqrt{K} a \cos \gamma \sin \gamma}{\cos^2 \sqrt{K} a - \sin^2 \gamma}$$

gegeben werden.

Die Kurven (7.23) stellen im allgemeinen Kurven zweiten Grades dar. Für $K > 0$ zerfallen diese in eine Ellipsen- bzw. Hyperbelfamilie, je nachdem der Ausdruck $\cos^2 \sqrt{K} a - \sin^2 \gamma$ positiv oder negativ ist. Für $K < 0$ stellt (7.27) eine Ellipsenfamilie dar. Alle diese Ellipsen gehen durch die Punkte (7.12). Die Gleichung des Horizykels erhält man aus (7.27), wenn man dort $\tanh \sqrt{-K} a = \cos \gamma$ setzt. Man findet dann als Gleichung dieser beiden zur u -Achse symmetrisch gelegenen Horizyklen die Gleichung

$$(7.26) \quad (1 \pm \sqrt{-K} Y)^2 = \operatorname{ch}^2 \sqrt{-K} a (1 + K(X^2 + Y^2)),$$

welche geometrisch zwei Ellipsen darstellen, die durch (7.12) gehen, und den Fundamentalkreis des Schemas in

$$(7.27) \quad X = 0, \quad Y = \pm \frac{1}{\sqrt{-K}}$$

berühren.

Da die Transformation (7.2) für die zugrunde gelegte Metrik gilt, so bedeutet hier wieder γ den Winkel, den die äußere Normale der linsenförmigen Extremale C_γ in den Punkten (7.12) mit der X -Achse bildet.

Für $Y_0 = 0$, also $\gamma = 0$, erhalten wir diejenige Extremale, die wir in den vorigen Nummern mit C_0 bezeichnet hatten. Es folgt dann aus (7.26)

$$(7.28) \quad X^2 + Y^2 = \frac{\operatorname{tg}^2 \sqrt{K} a}{K};$$

C_0 stellt also auf der Cayley-Ebene einen durch (7.12) gehenden Kreis mit dem Radius $\frac{\operatorname{tg} \sqrt{K} a}{\sqrt{K}}$ dar.

Damit ist das Wesentlichste über die linsenförmigen Extremalen C_γ gesagt worden. Um die Gleichung der Kreis-Rechteck-Extremalen aufzustellen, müssen wir vorher einiges über die Cayleyschen Bewegungen längs der Y -Achse wiederholen.

8. Die Gleichung der Kreis-Rechteck-Extremalen C_q^* . Die Gruppe, bei welcher die Kurve

$$(8.1) \quad X^2 - \operatorname{tg}^2 \sqrt{K} a Y^2 = \frac{\operatorname{tg}^2 \sqrt{K} a}{K}$$

in sich übergeht, bildet eine Untergruppe der allgemeinen Bewegungsgruppe der Cayley-Ebene, die Untergruppe der „Verschiebungen“ längs der Y -Achse.

Führt man die homogenen Koordinaten x_1, x_2, x_3 von Weierstraß ein, so lautet bekanntlich²¹⁾ die allgemeinste Bewegung

$$(8.2) \quad \begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3, \\ x_2' &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3, \\ x_3' &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{aligned}$$

mit der Determinante Eins.

Da nun durch Einführung der homogenen Koordinaten (8. 1) in die beiden Geraden

$$(8.3) \quad x_1^2 = \frac{\sin^2 \sqrt{K} a}{K}$$

übergeht, so folgt daraus unmittelbar

$$(8.4) \quad a_{11} = 1, \quad a_{12} = a_{22} = 0.$$

Ferner folgt aus

$$(8.5) \quad x_1'^2 + x_2'^2 + \frac{x_3'^2}{K} = x_1^2 + x_2^2 + \frac{x_3^2}{K} = \frac{1}{K}$$

unmittelbar

$$a_{21} = a_{31} = 0.$$

Die fragliche Untergruppe sieht also so aus

$$(8.6) \quad \begin{aligned} x_1' &= x_1, \\ x_2' &= a x_2 + b x_3, \\ x_3' &= c x_2 + d x_3 \end{aligned}$$

mit a, b, c, d Konstanten. Diese Konstanten erfüllen noch wegen (8. 5) die Bedingungen

$$(8.7) \quad a^2 + \frac{c^2}{K} = 1, \quad ab + \frac{cd}{K} = 0, \quad b^2 + \frac{d^2}{K} = \frac{1}{K},$$

und man kann sie mittels eines Parameters λ in folgender Form ausdrücken:

$$(8.8) \quad \begin{aligned} a &= \cos \sqrt{K} \lambda, \quad b = -\frac{\sin \sqrt{K} \lambda}{\sqrt{K}}, \quad c = \sqrt{K} \sin \sqrt{K} \lambda \\ &\text{und} \quad d = \cos \sqrt{K} \lambda. \end{aligned}$$

Es wird also mit Rücksicht auf (8. 7)

$$(8.9) \quad \begin{aligned} x_1' &= x_1, \\ x_2' &= x_2 \cos \sqrt{K} \lambda - x_3 \frac{\sin \sqrt{K} \lambda}{\sqrt{K}}, \\ x_3' &= x_3 \sqrt{K} \sin \sqrt{K} \lambda + x_2 \cos \sqrt{K} \lambda, \end{aligned}$$

²¹⁾ Vgl. z. B. Cartan, S. 170.

und diese Bewegung ist in der uv -Ebene gleichwertig mit der Bewegung

$$(8.10) \quad u' = u, \quad v' = v + \lambda.$$

Die Gleichung einer Kreis-Rechteck-Extremale vom Typus C_q^* ist jetzt

$$(8.11) \quad \sin^2 \sqrt{K} u + \cos^2 \sqrt{K} u \sin^2 \sqrt{K} \left(v \pm \frac{q}{2} \right) = \sin^2 \sqrt{K} a$$

für $-a < u < a$ und

$$(8.12) \quad -\frac{q}{2} \leq v \leq \frac{q}{2}$$

für $u = \pm a$.

Bezeichnet man also mit x_1^0, x_2^0, x_3^0 die homogenen Koordinaten der Mittelpunkte der beiden Halbkreise, aus denen C_q^* zum Teil gebildet ist, so kann man an Stelle von (8.11) schreiben

$$(8.13) \quad x_1^2 + (x_2 x_3^0 - x_3 x_2^0)^2 = \frac{\sin^2 \sqrt{K} a}{K}.$$

Daraus folgt mit Hilfe der früher entwickelten Transformationsformeln

$$(8.14) \quad X^2 + \frac{(Y \mp Y_0)^2}{1 + K Y_0^2} = \frac{\sin^2 \sqrt{K} a}{K} (1 + K(X^2 + Y^2)) \quad (|Y| \geq |Y_0|)$$

mit

$$(8.15) \quad Y_0 = \frac{\operatorname{tg} \sqrt{K} \frac{q}{2}}{\sqrt{K}}.$$

Die Gleichung (8.14) stellt zwei zur Y -Achse symmetrisch gelegene Halbellipsen dar, welche die Kurve

$$(8.16) \quad X^2 - X^2 \operatorname{tg}^2 \sqrt{K} a = \frac{\operatorname{tg}^2 \sqrt{K} a}{\sqrt{K}}$$

in den Punkten

$$(8.17) \quad \begin{aligned} X &= \pm \frac{1}{\sqrt{K}} \frac{\operatorname{tg} \sqrt{K} a}{\cos \sqrt{K} \frac{q}{2}}, \\ Y &= \pm \frac{1}{\sqrt{K}} \operatorname{tg} \sqrt{K} \frac{q}{2} \end{aligned}$$

berühren. Diese Kurvenstücke zusammen mit den entsprechenden Bögen von (8.16) begrenzen dann die Kreis-Rechteck-Extremale vom Typus C_q^* .

Somit ist auch dieser Fall erledigt.

Es erübrigt sich vollständig, den Fundamentalsatz für die Cayley-Ebene zu geben. Sein Wortlaut ist identisch mit demjenigen von Nr. 6, der einzige Unterschied ist nur, daß an Stelle von (5.9), (5.10), (5.12) die allgemeine Gleichung (7.23) eintritt. Entsprechend lautet die Änderung für die Kreis-Rechteck-Extremalen.

9. Die allgemeine isoperimetrische Aufgabe für die ebenen Geometrien konstanter Krümmung. Wir entwickeln zuerst einige neue Verschärfungen der isoperimetrischen Ungleichung für alle vier Geometrien konstanter Krümmung, welche schon längst analoge Verschärfungen in der ebenen euklidischen Geometrie besitzen. Dazu gehen wir von der Gleichung (6. 1) aus und lassen dort $\gamma \rightarrow 0$ streben. Wir erhalten dann

$$(9.1) \quad F - F_0 = \varrho_0(L - L_0) - 2 \varrho_0 R(C, C_0)$$

mit

$$(9.2) \quad \varrho_0 = \frac{\operatorname{tg} \sqrt{K} a}{\sqrt{K}}.$$

Nun ist die Extremale C_0 ein gewöhnlicher Breitenkreis mit Radius a und wir haben

$$(9.3) \quad \begin{aligned} F_0 &= 4\pi \frac{\sin^3 \frac{\sqrt{K} a}{2}}{K}, \\ L_0 &= 2\pi \frac{\sin \sqrt{K} a}{\sqrt{K}}. \end{aligned}$$

Hiermit wird nach einigen Rechnungen

$$(9.4) \quad F = \frac{\operatorname{tg} \sqrt{K} a}{\sqrt{K}} L + \frac{2\pi}{K} \left(1 - \frac{1}{\cos \sqrt{K} a}\right) - \frac{2 \operatorname{tg} \sqrt{K} a}{\sqrt{K}} R(C, C_0)$$

oder

$$(9.5) \quad \frac{2\pi}{K} = \frac{\sin \sqrt{K} a}{\sqrt{K}} L + \frac{\cos \sqrt{K} a}{K} (2\pi - KF) - 2 \frac{\sin \sqrt{K} a}{K} R(C, C_0).$$

Ist $K \leq 0$, so ist die Größe

$$(9.6) \quad Q = 2\pi - KF$$

positiv. Ist aber $K > 0$ und liegt daher der elliptische Fall vor, so ist, da der Gesamthalt der elliptischen Ebene $\frac{2\pi}{\sqrt{K}}$ ist, ebenfalls $Q \geq 0$. Hiermit ist die Ungleichung $Q \geq 0$ für alle drei Geometrien bewiesen. Da nun für den sphärischen Fall ebenfalls²²⁾

$$(9.7) \quad F < \frac{2\pi}{\sqrt{K}}$$

angenommen werden kann, so dürfen wir im folgenden Q nicht negativ voraussetzen.

Nun folgt aus (9. 5)

$$(9.8) \quad 2\pi = \sqrt{K} \sin \sqrt{K} a L + \cos \sqrt{K} a Q - 2 \sqrt{K} \sin \sqrt{K} a R(C, C_0),$$

²²⁾ Vgl. z. B. die Arbeit von Erhard Schmidt (Schmidt III, S. 752).

und daraus ergibt sich für $K > 0$ wegen $R = R(C, C_0) \geq 0$

$$(9.9) \quad 2\pi \leq \sqrt{K} \sin \sqrt{K} a L + \cos \sqrt{K} a Q,$$

und für $K < 0$

$$(9.10) \quad 2\pi \geq \sqrt{K} \sin \sqrt{K} a L + \cos \sqrt{K} a Q.$$

Zur Erhaltung einer allgemeinen isoperimetrischen Ungleichung setzen wir für den Augenblick

$$(9.11) \quad \Phi(a) = \sqrt{K} \sin \sqrt{K} a L + \cos \sqrt{K} a Q.$$

Dann wird

$$(9.12) \quad \Phi^2(a) = K L^2 + Q^2 - K \cos^2 \sqrt{K} a \left(L - Q \frac{\operatorname{tg} \sqrt{K} a}{\sqrt{K}} \right)^2,$$

also für $K > 0$ mit Rücksicht auf (9.9)

$$(9.13) \quad 4\pi^2 \leq K L^2 + Q^2 - K \cos^2 \sqrt{K} a \left(L - Q \frac{\operatorname{tg} \sqrt{K} a}{\sqrt{K}} \right)^2.$$

Daraus folgt wegen $Q = 2\pi - KF$ die Ungleichung

$$(9.14) \quad F(4\pi - KF) \leq L^2 - \cos^2 \sqrt{K} a \left(L - Q \frac{\operatorname{tg} \sqrt{K} a}{\sqrt{K}} \right)^2.$$

Ist $K < 0$, so hat $\Phi(a)$ ein einziges Minimum für

$$(9.15) \quad \frac{\operatorname{tg} \sqrt{K} a}{\sqrt{K}} = \frac{L}{Q}, \quad \text{also für } a = \frac{1}{\sqrt{K}} \arctg \frac{\sqrt{K} L}{Q}.$$

Eine leichte Rechnung zeigt dann, daß für diesen Wert von a , $\Phi \geq 0$ ist. Wir können also (9.10) quadrieren und wir erhalten

$$(9.16) \quad 4\pi^2 \geq \Phi^2(a).$$

Daraus folgt mit Rücksicht auf (9.12) die Ungleichung

$$(9.17) \quad 4\pi^2 \geq K L^2 + Q^2 - K \cos^2 \sqrt{K} a \left(L - Q \frac{\operatorname{tg} \sqrt{K} a}{\sqrt{K}} \right)^2.$$

Ersetzt man jetzt Q durch $2\pi - KF$, so folgt nach einigen leichten Rechnungen dieselbe Ungleichung (9.14) wieder. Diese Ungleichung gilt also sowohl für $K > 0$ wie für $K < 0$ und geht für $K \rightarrow 0$ in die bekannte isoperimetrische Ungleichung

$$(9.18) \quad 4\pi F \leq L^2 - (L - 2\pi a)^2$$

über.

Die Ungleichung (9.14) verschärft mit Hilfe der Breite a die Ungleichung

$$(9.19) \quad F(4\pi - KF) \leq L^2,$$

die für $K < 0$ durch ähnliche Überlegungen zuerst von Erhard Schmidt²³⁾ bewiesen wurde. Das Gleichheitszeichen tritt in (9.14) nur dann ein, wenn $R(C, C_0) = 0$ ist, also wenn C bis auf Parallelverschiebung gleich C_0^* ist, d. h. wenn C eine Kreis-Rechteck-Extremale ist.

Ich füge noch einige Bemerkungen zu der Verschärfung (9.14) hinzu. Die Ungleichung (9.14) wurde abgeleitet unter der Voraussetzung, daß für $K > 0$ ein $a < \frac{\pi}{2\sqrt{K}}$ gibt, so daß C im Streifen

$$(9.20) \quad -a \leq u \leq a$$

eingeschrieben werden kann. Nun hat Radò bewiesen²⁴⁾, daß jede einfach geschlossene Kurve C auf einer Kugel mit dem Radius $R = \frac{1}{\sqrt{K}}$ von der Länge $L < 2\pi R$ durch eine offene Halbkugel derselben überdeckt werden kann.

Mit anderen Worten:

Ist C eine einfach geschlossene Kurve auf der Kugel \mathfrak{R} .

$$(9.21) \quad x^2 + y^2 + z^2 = \frac{1}{K},$$

von der Länge $L < 2\pi R$, so existiert immer eine offene Halbkugel von \mathfrak{R} , welche C überdeckt.

Für den Beweis dieses Lemmas verweisen wir auf die Originalarbeit von Radò.

Dieses Lemma sichert nun im sphärischen Falle die Existenz eines Streifens von der Breite $2a < \frac{\pi}{\sqrt{K}}$, in dem C eingeschrieben werden kann.

Da wir uns ferner für den elliptischen Fall nur auf Kurven beschränken, die durch eine „Bewegung“ ins Endliche gebracht werden (also im sphärischen Abbild der Cayley-Ebene auf eine Halbkugel), so ist auch für diesen Fall die Bedingung erfüllt und somit für $K > 0$ (elliptisch und sphärisch) die Ungleichung

$$(9.22) \quad F(4\pi - KF) \leq L^2 - \cos^2 \sqrt{K} a \left(L - Q \frac{\operatorname{tg} \sqrt{K} a}{\sqrt{K}} \right)^2$$

für alle Kurven C bewiesen, deren Länge kleiner als $\frac{2\pi}{\sqrt{K}}$ ist. Für $K \leq 0$ gilt

(9.22) ohne die Einschränkung über die Länge.

Zum Schluß dieser Nummer einige Worte über die Verallgemeinerung einer bekannten Ungleichung von Bonnesen.

²³⁾ Schmidt III, S. 754 unten. Die unschärfere Ungleichung $4\pi F \leq L^2$ hatten schon früher Bechenbach und Radò [Trans. Amer. Mathem. Soc. 35 (1933), S. 662–674] bewiesen.

²⁴⁾ Tibor Radò, The isoperimetric inequality on the sphere. Amer. Journ. of Math. 57 (1935), S. 765–770.

Es bezeichne $2a_0$ das Maximum der „Breite“ von C für alle möglichen Richtungen. Wegen $L < \frac{2\pi}{\sqrt{K}}$ ist sowohl für $K \leq 0$ wie für $K > 0$

$$(9.23) \quad D = \cos \sqrt{K} a_0 > 0$$

und wir erhalten aus (9.22) für $K > 0$

$$(9.24) \quad F(4\pi - KF) \leq L^2 - D^2 \left(L - Q \frac{\operatorname{tg} \sqrt{K} a}{\sqrt{K}} \right)^2,$$

für jedes a . Schreibt man diese Ungleichung für zwei Breiten a_1 und a_2 , addiert man die so erhaltenen Resultate, so erhält man unter Verwendung der Ungleichung

$$(9.25) \quad \left(\frac{x+y}{2} \right)^2 \leq \frac{x^2 + y^2}{2}$$

die Ungleichung

$$(9.26) \quad F(4\pi - KF) \leq L^2 - \frac{D^2 (2\pi - KF)^2}{4} \left| \frac{\operatorname{tg} \sqrt{K} a_1}{\sqrt{K}} - \frac{\operatorname{tg} \sqrt{K} a_2}{\sqrt{K}} \right|^2.$$

Diese Ungleichung geht für $K \rightarrow 0$ (wegen $D \rightarrow 1$) in die bekannte Bonnesen'sche Ungleichung

$$(9.27) \quad 4\pi F \leq L^2 - \pi^2 |a_1 - a_2|^2$$

über.

Ist $K < 0$, so folgt aus (9.22) wegen $\cos \sqrt{K} a \geq 1$,

$$(9.28) \quad F(4\pi - KF) \leq L^2 - \left(L - Q \frac{\operatorname{tg} \sqrt{K} a}{\sqrt{K}} \right)^2$$

und daraus durch ähnliche Überlegungen wie vorhin und mit Rücksicht auf $Q \geq 2\pi$

$$(9.29) \quad F(4\pi - KF) \leq L^2 - \pi^2 \left| \frac{\operatorname{tg} \sqrt{K} a_1}{\sqrt{K}} - \frac{\operatorname{tg} \sqrt{K} a_2}{\sqrt{K}} \right|^2.$$

Diese Ungleichung geht für $K \rightarrow 0$ ebenfalls in die (9.27) über.

Daß aus den verschärften Ungleichungen die klassische isoperimetrische Eigenschaft des Kreises folgt, ist trivial und bedarf wohl keiner besonderen Erwähnung.

Kapitel III.

Die klassische isoperimetrische Aufgabe in einem n -dimensionalen Raum konstanter Krümmung.

10. *Elementare Behandlung einer allgemeinen Aufgabe von Erhard Schmidt für einen speziellen Fall.* Wir betrachten ähnlich wie in Nr. 2 die dort zugrunde gelegte Mannigfaltigkeit \mathfrak{M} mit der Metrik

$$(10.1) \quad ds^2 = du^2 + \varphi^2(u) dv^2$$

und machen hinsichtlich $\varphi(u)$ dieselben Voraussetzungen wie dort.

Wir behandeln zuerst folgende Aufgabe²⁵⁾:

Unter allen einfach in bezug auf $u = 0$ symmetrischen, geschlossenen Kurven C , welche im Streifen

$$(10.2) \quad -a \leq u \leq a$$

einbeschrieben sind, soll diejenige bestimmt werden, die bei konstantem (für alle C vorgegebenem) Integral

$$(10.3) \quad G = \int_{C_1} \psi^{n-2}(u) ds \quad (n \text{ ganz } \geq 2)$$

das Maximum von

$$(10.4) \quad J = \int_{C_1} \psi^{n-1}(u) \frac{dv}{ds} ds$$

liefert. Dabei ist $\psi(u) \geq 0$ die schon in Nr. 2 eingeführte Funktion

$$(10.5) \quad \psi(u) = \int_0^u \varphi(x) dx.$$

In beiden Integralen soll die Integration über die rechte Hälfte C_1 von C erstreckt werden ($u \geq 0$), und zwar vom untersten Punkt P_0 ($u = 0, v = v(0)$) bis zum obersten Punkt P_1 ($u = 0, v = v(l)$). l ist hier die Gesamtlänge von C_1 . Es wird noch vorausgesetzt, daß C_1 mit der Symmetrieachse $u = 0$ nur diese beiden Punkte gemeinsam hat.

Nach den Voraussetzungen von Nr. 2 ist die Funktion

$$(10.6) \quad \omega(u) = \frac{\psi(u)}{\varphi(u)}$$

im Intervall (3.2) monoton wachsend und es gilt $\omega'(a) > 0$.

Wir behaupten jetzt, daß die Extremalen unseres Problems mit den Extremalen des Problems von Nr. 2 übereinstimmen, also wieder Bögen von geodätischen Kreisen sind, die nach Spiegelung auf $u = 0$ durch die Punkte ($u = \pm a, v = 0$) gehen. In der Tat konstruieren wir wie in Nr. 2 die Extremalfamilien F_1 und F_2 . Jede Kurve von F_1 möge der Differentialgleichung

$$(10.7) \quad \frac{dv}{du} + \frac{\omega(u)}{\varphi \sqrt{\varphi^2 - \omega^2(u)}} = 0$$

²⁵⁾ Bei Erhard Schmidt, l. c. V, treten an Stelle der Funktionen ψ^{n-2} , ψ^{n-1} willkürliche Funktionen $\sigma(u)$, $\tau(u)$ auf. An Stelle von (10.2) tritt ein beliebiger Streifen $a \leq u \leq b$ ein und Symmetrie der Kurven und der Metrik wird nicht verlangt. Auch die Monotonie von $\omega(u)$ wird nicht gefordert. Jedoch sind alle unsere Annahmen beim klassischen isoperimetrischen Problem für Riemannsche Räume konstanter Krümmung erfüllt. Da die vorliegende Arbeit nur die Behandlung dieses Problems bezweckt, so habe ich alle Nebenbetrachtungen beiseite gelassen, insbesondere diejenigen, welche auf mögliche Verallgemeinerungen der hier behandelten speziellen Fälle hinzielen.

genügen und entsprechend jede Kurve von F_2 der Differentialgleichung

$$(10.8) \quad \frac{dv}{du} - \frac{\omega(u)}{\varphi \sqrt{\varrho_\gamma^2 - \omega^2(u)}} = 0.$$

Die Kurven K_1, K_2 , welche durch $(\pm a, 0)$ gehen, werden dann durch die Gleichung

$$(10.9) \quad v \pm \int_{-a}^u \frac{\omega dx}{\varphi \sqrt{\varrho_\gamma^2 - \omega^2(x)}} = 0$$

gegeben. Das von ihnen begrenzte Gebiet (das den Nullpunkt enthält) bezeichnen wir wieder mit \mathfrak{R}_γ . ϱ_γ ist hier wieder gleich $\varrho_0: \cos \gamma$ ($\varrho_0 = \omega(a)$) und γ ist gleich dem Winkel, den die äußeren Normalen \bar{n} von \mathfrak{R}_γ in den Punkten $(\pm a, 0)$ mit der v -Linie bilden. Der Beweis der Behauptung kann leicht erbracht werden, wenn man genau wie in Nr. 2 die Punkte einer beliebigen Kurve C den Punkten einer „Extremalen“ C_γ zuordnet.

Man erhält wieder mit Rücksicht auf (2.11), (2.12)

$$(10.10) \quad A = \int_{C_1} \cos(n, \bar{n}) \varphi^{n-2}(u) ds = k_\gamma \int_{C_1} \varphi^{n-2} \omega \varphi \frac{dv}{ds} ds + \\ + \int_{C_1} \varphi^{n-2} \sqrt{1 - K_\gamma^2 \omega^2} \frac{du}{ds} ds.$$

Daraus folgt

$$(10.11) \quad A = k_\gamma \int_{C_1} \varphi^{n-1}(u) \frac{dv}{ds} ds + \int_{C_1} \varphi^{n-2}(u) \sqrt{1 - k_\gamma^2 \omega^2} \frac{du}{ds} ds,$$

oder noch durch genau ähnliche Überlegungen wie in Nr. 2

$$(10.12) \quad A = k_\gamma J + 2 \int_0^a \varphi^{n-2}(u) \sqrt{1 - k_\gamma^2 \omega^2(u)} du.$$

Durch Spezialisierung $C = C_\gamma$, Subtraktion der so entstandenen Gleichungen und Ausführung einiger elementarer Rechnungen ähnlich wie in Nr. 2 erhält man schließlich

$$(10.13) \quad J - J_\gamma = \varrho_\gamma(G - G_\gamma) - 2 \varrho_\gamma R_n(C, C_\gamma).$$

Dabei sind J_γ, G_γ die Werte von (10.3), (10.4) für die Extremale C_γ , $\varrho_\gamma = \frac{1}{k_\gamma}$ und $R_n(C, C_\gamma)$ stellt das Restglied

$$(10.14) \quad R_n(C, C_\gamma) = 2 \int_{C_1} \varphi^{n-2}(u) \sin^2 \frac{(n, \bar{n})}{2} ds$$

dar. Es braucht hier kaum wiederholt zu werden, daß man aus (10.13) durch genau dieselben Überlegungen wie in Nr. 2 die entsprechende Gleichung

für die „Kreis-Rechteck-Extremalen“ C_q^* gewonnen hat. Wir verzichten auf eine ausführliche Wiedergabe dieser Gleichung, zumal der Leser in Nr. 2 alles Nötige findet, um diese Gleichung zu erhalten.

Wichtig ist es zu bemerken, daß durch den Wert von G , C_γ bzw. G_q^* eindeutig bestimmt wird. Denn rechnet man die Bogenlänge von C_γ aus, so folgt mit Rücksicht auf C (10. 1), (10. 7), (10. 8)

$$(10. 15) \quad ds = \frac{|du|}{\sqrt{1 - k_\gamma^2 \omega^2}} \quad (k_\gamma = \frac{1}{e_\gamma});$$

es ist also

$$(10. 16) \quad G = G_\gamma = 2 \int_0^a \frac{\psi^{n-2} du}{\sqrt{1 - k_\gamma^2 \omega^2}},$$

woraus die Monotonie von G_γ als Funktion von γ unmittelbar folgt. Bezeichnet man ferner mit G_0 das Integral

$$(10. 17) \quad G_0 = 2 \int_0^a \frac{\psi^{n-2} du}{\sqrt{1 - k_0^2 \omega^2}},$$

so gilt für den Wert G_q^* von C_q^*

$$(10. 18) \quad G_q^* = G_0 + 2 \int_0^a \varphi(a) \psi^{n-2}(a).$$

Hat man also einen Wert G von (10. 3) mit

$$(10. 19) \quad G > 2 \int_0^a \psi^{n-2}(u) du$$

angegeben, so entspricht jedem G eine und nur eine Extremale und diese ist nun „linsenförmig“ bzw. vom „Kreis-Rechteck-Typus“, je nachdem G kleiner G_0 oder größer gleich G_0 ist. Dies alles gilt natürlich nur dann unbeschränkt, wenn man die Schlichtheit der Bedeckung außer acht läßt. Sonst treten in (10. 18) Einschränkungen nach oben ein, die wir später am richtigen Platz ausführlich aufstellen werden.

11. *Die räumliche isoperimetrische Aufgabe. Vorbereitende Tatsachen.* Bevor wir zur Aufstellung und Lösung der räumlichen Aufgabe übergehen, wiederholen wir kurz die fundamentalen Tatsachen der räumlichen Cayleyschen Geometrie. Für ein tieferes Eindringen in diese verweisen wir auf das erwähnte Buch von Cartan.

Das Cayleysche Schema der nicht-euklidischen Geometrie im Raume verwendet analog wie im Falle der Ebene eine Fläche zweiten Grades Q als „Fundamentalfäche“.

Man unterscheidet je nach der Natur von Q drei Fälle:

1. Q ist imaginär mit reeller Gleichung (Pseudoreell). Dann gehören alle Punkte von R_n dem „Gebäude“ an und die Geometrie erstreckt sich auf alle Punkte von R_n . Man kann ohne Einschränkung der Allgemeinheit als Gleichung von Q die Gleichung

$$(11.1) \quad X_1^2 + X_2^2 + \dots + X_n^2 + \frac{1}{K} = 0 \quad (K > 0)$$

setzen. Dabei sind X_i die Koordinaten eines Punktes von R_n in bezug auf ein rechtwinkliges Koordinatensystem O .

2. Q ist eine reelle, konvexe, geschlossene Fläche. Die Geometrie erstreckt sich nur auf das Innere von Q . Man kann ohne weiteres als Q die Kugel

$$(11.2) \quad X_1^2 + X_2^2 + \dots + X_n^2 + \frac{1}{K} = 0$$

wählen.

3. $K = 0$. Die Gleichung der Fundamentalfläche lautet

$$(11.3) \quad X_1^2 + X_2^2 + \dots + X_n^2 = 0.$$

Das ist der Fall der gewöhnlichen euklidischen Geometrie.

Sind nun $P(X_i)$, $P'(X_i + dX_i)$ ($i = 1, 2, \dots, n$) zwei benachbarte Punkte, so definiert man die Entfernung dieser Punkte durch

$$(11.4) \quad ds^2 = \frac{(1 + K \sum_1^n K_i^2) (\sum dX_i^2) - K (\sum X_i dX_i)^2}{(1 + K \sum X_i^2)^2},$$

wobei $K \geq 0$ ist.

Nach diesen vorbereitenden Tatsachen führen wir neue „krummlinige“ Koordinaten ein, und zwar durch die Formeln

$$(11.5) \quad \begin{aligned} X_1 &= \frac{1}{\sqrt{K}} \frac{\operatorname{tg} \sqrt{K} \vartheta_1}{\cos \sqrt{K} \vartheta_2 \dots \cos \sqrt{K} \vartheta_n}, \\ X_2 &= \frac{1}{\sqrt{K}} \frac{\operatorname{tg} \sqrt{K} \vartheta_2}{\cos \sqrt{K} \vartheta_3 \dots \cos \sqrt{K} \vartheta_n}, \\ &\dots \dots \dots \\ X_{n-1} &= \frac{1}{\sqrt{K}} \frac{\operatorname{tg} \sqrt{K} \vartheta_{n-1}}{\cos \sqrt{K} \vartheta_n}, \\ X_n &= \frac{1}{\sqrt{K}} \operatorname{tg} \sqrt{K} \vartheta_n, \end{aligned}$$

und berechnen ds^2 in diesen neuen Koordinaten.

Die ϑ_i variieren hier für $K > 0$ in den Intervallen

$$(11.6) \quad -\frac{\pi}{2\sqrt{K}} \leq \vartheta \leq \frac{\pi}{2\sqrt{K}},$$

Zuerst folgt aus (11. 8), (11. 9)

$$(11. 12) \quad 1 + K \sum_1^n X_i^2 = \frac{1}{x_{n+1}^2}.$$

Es wird also

$$(11. 13) \quad ds^2 = x_n^2 \left(\sum_1^n dX_i^2 \right) - K \left(\sum_1^n X_i dX_i \right).$$

Nun ist mit Rücksicht auf (11. 12)

$$(11. 14) \quad K \left(\sum_1^n X_i dX_i \right) = x_{n+1}^{-3} dx_{n+1}$$

und man erhält durch Einsetzen in (11. 13) mit Rücksicht auf (11. 9)

$$(11. 15) \quad x_{n+1}^2 dX_i = x_{n+1} dx_i - x_i dx_{n+1}.$$

Somit wird

$$(11. 16) \quad ds^2 = \frac{\sum_1^n (x_{n+1} dx_i - x_i dx_{n+1})^2}{x_{n+1}^2} - \frac{1}{K x_{n+1}^2} dx_{n+1}^2,$$

also

$$(11. 17) \quad ds^2 = \sum_1^n dx_i^2 + \frac{1}{K} dx_{n+1}^2 - \frac{2 dx_{n+1}}{x_{n+1}} \left(\sum_1^n x_i dx_i + \frac{1}{K} x_{n+1} dx_{n+1} \right).$$

Da nun wegen (11. 8) die letzte Klammer Null ist, so folgt daraus (11. 11) unmittelbar.

Im elliptischen Fall $K > 0$ nennen wir ähnlich wie für $n = 2$ den Bereich

$$(11. 18) \quad -\frac{\pi}{2\sqrt{K}} \leq \vartheta_i < \frac{\pi}{2\sqrt{K}} \quad (i = 1, 2, \dots, n)$$

den Fundamentalbereich des elliptischen Raumes und bezeichnen ihn mit G_0 .

Die Bewegungen, die uns hier interessieren, sind diejenigen Bewegungen des elliptischen Raumes, die (1. 4) in sich überführen, und zwar ohne Rotation derselben um die X_n -Achse. Solche Bewegungen werden durch

$$(11. 19) \quad \begin{aligned} x_i' &= x_i & (i = 1, 2, \dots, n-1), \\ x_n' &= ax_n + bx_{n+1}, \\ x_{n+1}' &= cx_n + dx_{n+1} \end{aligned}$$

gegeben und man rechnet genau wie im ebenen Fall $n = 2$ um, daß dabei die ϑ_i ($i = 1, 2, \dots, n$) folgende Transformation erleiden

$$(11. 20) \quad \begin{aligned} \vartheta_i' &= \vartheta_i & (i = 1, 2, \dots, n-1), \\ \vartheta_n' &= \vartheta_n + \lambda \end{aligned}$$

mit einem willkürlichen Parameter λ .

Wir werden hier darauf nicht näher eingehen, zumal wir den Körper \mathfrak{R} von vornherein innerhalb des Fundamentalbereiches G_0 annehmen werden.

$$(11.21) \quad x_1^2 + x_2^2 + x_{n+1}^2 = \frac{1}{K} \quad (K > 0)$$

$$\begin{aligned}
 x_1 &= \frac{1}{\sqrt{K}} \sin \sqrt{K} \vartheta_1, \\
 x_2 &= \frac{1}{\sqrt{K}} \cos \sqrt{K} \vartheta_1 \sin \sqrt{K} \vartheta_2, \\
 &\dots \dots \dots \\
 x_n &= \frac{1}{\sqrt{K}} \cos \sqrt{K} \vartheta_1 \cos \sqrt{K} \vartheta_2 \dots \sin \sqrt{K} \vartheta_n, \\
 x_{n+1} &= \frac{1}{\sqrt{K}} \cos \sqrt{K} \vartheta_1 \cos \sqrt{K} \vartheta_2 \dots \cos \sqrt{K} \vartheta_n.
 \end{aligned}
 \tag{11.22}$$

$$(11.23) \quad -\frac{\pi}{2\sqrt{K}} < \theta_i < \frac{\pi}{2\sqrt{K}} \quad (i = 1, 2, \dots, n-1), \quad -\frac{\pi}{\sqrt{K}} \leq \theta_n < \frac{\pi}{\sqrt{K}}$$

$$(12.1) \quad X_1^2 + X_2^2 + \dots + X_{n-1}^2 - X_n^2 \operatorname{tg}^2 \sqrt{K} a = \frac{\operatorname{tg}^2 \sqrt{K} a}{K}$$

$$(12.2) \quad X_1^2 + X_2^2 + \dots + X_{n-1}^2 - X_n^2 \operatorname{tg}^2 \sqrt{K} a \leq \frac{\operatorname{tg}^2 \sqrt{K} a}{K}$$

$$(12.3) \quad x_1^2 + x_2^2 + \dots + x_{n-1}^2 \leq \frac{\sin^2 \sqrt{K} a}{K}$$

einbeschrieben sind und die eine vorgegebene Oberfläche O besitzen, soll derjenige bestimmt werden, der das Maximum des eingeschlossenen Gebietes aufweist.

In dem Raum Θ_n der ϑ_i -Koordinaten hat dieselbe Aufgabe folgenden Wortlaut:

Unter allen Rotationskörpern \mathfrak{R} , welche im Fundamentalbereich des Raumes Θ_n mit der Metrik

$$(12.4) \quad ds^2 = d\vartheta_1^2 + \cos^2 \sqrt{K} \vartheta_1 d\vartheta_2^2 + \dots + \\ + \cos^2 \sqrt{K} \vartheta_1 \dots \cos^2 \sqrt{K} \vartheta_{n-1} d\vartheta_n^2$$

liegen und im Zylinder \mathfrak{S}_a

$$(12.5) \quad \frac{\cos^2 \sqrt{K} \vartheta_1 \dots \cos^2 \sqrt{K} \vartheta_{n-1}}{K} \geq \frac{\cos^2 \sqrt{K} a}{K}$$

einbeschrieben sind, soll bei vorgegebener Oberfläche O derjenige Körper bestimmt werden, der das Maximum des eingeschlossenen Volumens aufweist.

Für $K > 0$ sind die Fundamentalbereiche des elliptischen bzw. sphärischen Raumes angegeben worden. Alle Rotationskörper, die durch eine „Bewegung“ parallel zu ϑ_n Platz in diesem finden können, werden von unserem Standpunkt als zulässige Körper betrachtet. Für $K \leq 0$ umfaßt der Variabilitätsbereich der ϑ_i den ganzen euklidischen Raum R_n und der Zylinder \mathfrak{S}_a ist ein in ihm offener Zylinder.

Für $K = 0$, der hier immer als Grenzfall betrachtet werden muß, muß man in (12.4) $\cos^2 \sqrt{K} a$ bzw. die $\cos^2 \sqrt{K} \vartheta_i$ durch $1 - \sin^2 \sqrt{K} a$ bzw. $1 - \sin^2 \sqrt{K} \vartheta_i$ ersetzen und dann K gegen Null streben lassen. Man erhält an Stelle von (12.5) den gewöhnlichen Zylinder

$$(12.6) \quad \vartheta_1^2 + \vartheta_2^2 + \dots + \vartheta_{n-1}^2 \leq a^2,$$

der immer in diesem Fall gemeint werden soll.

Bevor wir jetzt zur Lösung der zuletzt gestellten Aufgabe übergehen, schicken wir einiges über die Bestimmung des Volumens bzw. der Oberfläche eines Rotationskörpers \mathfrak{R} in einem n -dimensionalen Riemannschen Raum konstanter Krümmung voraus

13. *Über das Volumen und die Oberfläche von Rotationskörpern in Θ_n .* Wir entwickeln in dieser Nummer die allgemeinen Formeln für den Inhalt und die Oberfläche eines Rotationskörpers in elliptisch-sphärischem bzw. hyperbolischem Raum. Die entsprechenden Formeln für den euklidischen Raum können dann leicht durch Grenzübergang erhalten werden.

Als Definition eines Rotationskörpers \mathfrak{R} in Θ_n wird folgende Definition zugrunde gelegt: Ein Körper \mathfrak{R} wird Rotationskörper heißen, falls er bei allen Drehungen des betreffenden Raumes, welche eine nicht euklidische Gerade ϱ

invariant lassen, in sich übergeht. Die Rotationen in Θ_n werden durch lineare Transformationen der homogenen Koordinaten

$$(13.1) \quad x'_i = \sum_1^{n+1} a_{ik} x_k \quad (i = 1, 2, \dots, n+1)$$

gekennzeichnet, welche die Form (11.8) und eine nichteuklidische Gerade invariant lassen²⁷⁾.

Im Falle des n -dimensionalen sphärischen Raumes handelt es sich hierbei um Drehungen der Kugel

$$(13.2) \quad x_1^2 + x_2^2 + \dots + x_{n+1}^2 = \frac{1}{K} \quad (K > 0),$$

die einen Großkreis derselben invariant lassen.

Nun wählen wir das Koordinatensystem so, daß die Rotationsachse mit der Parameterlinie θ_n übereinstimmt, also die Gleichung

$$(13.3) \quad \theta_1 = \theta_2 = \dots = \theta_{n-1} = 0$$

hat. Es wird noch ferner vorausgesetzt, daß die Oberfläche von \mathfrak{R} mit der Rotationsachse zwei und nur zwei verschiedene Punkte gemeinsam hat. In den θ_i kann dann die Oberfläche \mathfrak{D} von \mathfrak{R} durch die Gleichungen

$$(13.4) \quad \begin{aligned} \theta_n &= \theta_n(t), \\ \cos \sqrt{K} \theta_1 \cos \sqrt{K} \theta_2 \dots \cos \sqrt{K} \theta_{n-1} &= \cos \sqrt{K} u(t) \quad (\bar{n} \geq 0) \end{aligned}$$

dargestellt werden, wobei der Parameter t von t_1 bis t_2 variiert. Es wird noch $\theta_n(t_1) < \theta_n(t_2)$ vorausgesetzt.

Diese Darstellung ist eine unmittelbare Folgerung der Definitionsgleichung von \mathfrak{R}

$$(13.5) \quad x_n = x_n(t), \quad x_{n+1} = x_{n+1}(t), \quad \sum_1^{n-1} x_r^2 = r^2(t) = \frac{\sin^2 \sqrt{K} u(t)}{K},$$

wenn man noch die Gleichung

$$(13.6) \quad r^2 + x_n^2 + \frac{x_{n+1}^2}{K} = \frac{1}{K} \quad (28)$$

berücksichtigt. Die Bedeutung von $\sqrt{|K|} u(t)$ ist folgende:

Betrachtet man die Hyperebene $\theta_n = \theta_n(t)$, so ist $\sqrt{|K|} u(t)$ der sphärische Radius des Schnittes dieser Großkugel mit dem Rotationskörper \mathfrak{R} .

Um jetzt \mathfrak{R} durch Rotation einer Kurve C zu erzeugen, betrachten wir im dreidimensionalen Raum r, x_n, x_{n+1}

$$(13.7) \quad r^2 + x_n^2 + \frac{x_{n+1}^2}{K} = \frac{1}{K} \quad (29)$$

²⁷⁾ Vgl. darüber Cartan, l. c.

²⁸⁾ Vgl. hier die Ausführungen von Erhard Schmidt in III, S. 769.

²⁹⁾ Im Falle des sphärischen Raumes schreiben wir an Stelle von (13.7) $x_{n+1} = \sqrt{K} x'_{n+1}$.

konstanter Krümmung K die „Meridiankurve“ C , deren laufende Koordinaten durch

$$(13.8) \quad r = r(t) \geq 0, \quad x_n = x_n(t), \quad x_{n+1} = x_{n+1}(t), \quad (t_1 \leq t \leq t_2)$$

dargestellt werden. Es wird im folgenden vorausgesetzt, daß C im Fundamentalbereich liegt. Wir schreiben nun y_1, y_2, y_3 an Stelle von r, x_n, x_{n+1} und erfüllen (13.7) durch

$$(13.9) \quad \begin{aligned} y_1 &= \frac{1}{\sqrt{K}} \sin \sqrt{K} u, \\ y_2 &= \frac{1}{\sqrt{K}} \cos \sqrt{K} u \sin \sqrt{K} v, \\ y_3 &= \cos \sqrt{K} u \cos \sqrt{K} v. \end{aligned}$$

Auf C sind dann u und v Funktionen von t . Wir machen hier die Annahme, daß u', v' bis auf höchstens endlich viele Punkte bestimmt und stetig sind³⁰⁾.

Nach Voraussetzung ist in $t_1 \leq t \leq t_2$

$\tau \geq 0$

Das Bogenelement ds von C wird, wie man leicht nachrechnet, durch

(13.10) $ds = \sqrt{u'^2 + \cos^2 \sqrt{K} u v'^2} dt$
gegeben.

Um nun die Entfernung ds^* zweier benachbarter Punkte P, P' auf der Oberfläche \mathfrak{D} von \mathfrak{R} zu berechnen, versuchen wir, diese mit Hilfe von ds und dem Radius r des Schnittes (13.5), der P enthält, auszudrücken.

Wir behaupten: Es gilt

$$(13.11) \quad ds^{2*} = ds^2 + r^2 ds_K^2,$$

wobei $rd s_K$ das Bogenelement von \mathfrak{G} ,

$$(13.12) \quad \sum_1^{n-1} x_r^2 = r^2 = \frac{\sin^2 \sqrt{K} u}{K}$$

darstellt. Zum Beweis führen wir auf \mathfrak{S}_P Polarkoordinaten $\varphi_1, \varphi_2, \dots, \varphi_{n-2}$ ein und berücksichtigen noch die zwei letzten Gleichungen (13.9). Es wird

$$\begin{aligned}
 x_1 &= r \sin \varphi_1, \\
 x_2 &= r \cos \varphi_1 \sin \varphi_2, \\
 &\dots \dots \dots \\
 x_{n+1} &= r \cos \varphi_1 \cos \varphi_2 \dots \sin \varphi_{n-2}, \\
 x_n &= \frac{1}{\sqrt{K}} \cos \sqrt{K} u \sin \sqrt{K} v, \\
 x_{n+1} &= \cos \sqrt{K} u \cos \sqrt{K} v.
 \end{aligned}
 \tag{13.13}$$

³⁰⁾ An sich genügt es, u, v totalstetig vorauszusetzen.

Es folgt dann aus (11. 11)

$$(13. 14) \quad ds^2 = \sum_1^n dx_i^2 + \frac{1}{K} dx_{n+1}^2 = r^2 ds_K^2 + dr^2 + \sin^2 \sqrt{K} v du^2 + \\ + \cos^2 \sqrt{K} v dv^2.$$

Berücksichtigt man jetzt noch

$$(13. 15) \quad dr = \cos \sqrt{K} u du,$$

so folgt wegen (13. 10) die Behauptung (13. 11) unmittelbar. Nachdem (13. 11) bewiesen ist, so kann unmittelbar daraus geschlossen werden, daß das Flächenelement dO von \mathfrak{R} durch

$$(13. 16) \quad dO = r^{n-2} d\omega_{n-1} ds = \left(\frac{\sin \sqrt{K} u}{\sqrt{K}} \right)^{n-2} d\omega_{n-1} ds$$

gegeben wird. Hierin bedeutet $d\omega_{n-1}$ das Flächenelement der Einheitskugel \mathfrak{S}_{n-1}

$$(13. 17) \quad x_1^2 + x_2^2 + \dots + x_{n-1}^2 = 1.$$

Daraus folgt durch Integration über die ganze Oberfläche \mathfrak{O} von \mathfrak{R} bzw. über die Kurve C

$$(13. 18) \quad O = (n-1) W_{n-1} \int_C \left(\frac{\sin \sqrt{K} u}{\sqrt{K}} \right)^{n-2} ds,$$

wenn

$$(n-1) W_{n-1} = \int d\omega_{n-1}$$

die Gesamtoberfläche von \mathfrak{S}_{n-1} darstellt.

Um nun das Volumen V von \mathfrak{R} auszurechnen, greifen wir auf die Darstellung (11. 7) zurück. Danach ist

$$(13. 19) \quad V = \int \cos^{n-1} \sqrt{K} \vartheta_1 \cos^{n-2} \sqrt{K} \vartheta_2 \dots \cos \sqrt{K} \vartheta_{n-1} d\vartheta_1 \dots d\vartheta_n,$$

erstreckt über das Innere von \mathfrak{R} .

Wir rechnen zuerst das Integral

$$(13. 20) \quad J(u) = \int \cos^{n-1} \sqrt{K} \vartheta_1 \dots \cos^2 \sqrt{K} \vartheta_{n-1} d\vartheta_1 \dots d\vartheta_{n-1},$$

erstreckt auf die Mannigfaltigkeit

$$(13. 21) \quad \frac{\cos \sqrt{K} \vartheta_1 \cos \sqrt{K} \vartheta_2 \dots \cos \sqrt{K} \vartheta_{n-1}}{K} \geq \frac{\cos \sqrt{K} u}{K}.$$

Dieses Integral habe ich in meiner in 1) zitierten Arbeit²¹⁾ für $K = 1$ gerechnet. Ich gebe hier einen ausführlichen Beweis, der einheitlich für jedes K gilt. Dazu differenzieren wir $J(u)$ nach u . Es wird

$$(13.22) \quad J'(u) = \int \cos^{n-1} \sqrt{K} \vartheta_1 \dots \cos \sqrt{K} \vartheta_{n-1} d\vartheta_1 \dots d\vartheta_{n-1},$$

erstreckt über das Gebiet

$$(13.23) \quad \cos \sqrt{K} \vartheta_1 \dots \cos \sqrt{K} \vartheta_{n-1} = \cos \sqrt{K} u.$$

Hiermit wird aber

$$(13.24) \quad J'(u) = \cos \sqrt{K} u \int \cos^{n-2} \sqrt{K} \vartheta_1 \dots \cos \sqrt{K} \vartheta_{n-2} d\vartheta_1 \dots d\vartheta_{n-1} \\ = \cos \sqrt{K} u J_*(u).$$

Um $J_1(u)$ auszurechnen, setze man

$$\begin{aligned}
 x_1 &= \frac{1}{\sqrt{K}} \sin \sqrt{K} \vartheta_1, \\
 x_2 &= \frac{1}{\sqrt{K}} \cos \sqrt{K} \vartheta_1 \sin \sqrt{K} \vartheta_2, \\
 &\dots \dots \dots \\
 x_{n-1} &= \frac{1}{\sqrt{K}} \cos \sqrt{K} \vartheta_1 \cos \sqrt{K} \vartheta_2 \dots \sin \sqrt{K} \vartheta_{n-1}, \\
 y &= \cos \sqrt{K} \vartheta_1 \cos \sqrt{K} \vartheta_2 \dots \cos \sqrt{K} \vartheta_{n-1}.
 \end{aligned}
 \tag{13.25}$$

Es gilt dann

$$(13.26) \quad \sum_1^{n-1} x_r^2 + \frac{y^2}{K} = \frac{1}{K}$$

und $J_1(u)$ ist gleich der Oberfläche der geodätischen Kugel

$$(13.27) \quad \frac{y^2}{K} \geq \frac{\cos \sqrt{K} u}{K},$$

also nach bekannten Sätzen³²⁾

$$(13.28) \quad J_1(u) = (n-1) W_{n-1} \left(\frac{\sin \sqrt{K} u}{\sqrt{K}} \right)^{n-2}.$$

Hiermit wird

$$(13.29) \quad J'(u) = (u-1) W_{n-1} \cos \sqrt{K} u \left(\frac{\sin \sqrt{K} u}{\sqrt{K}} \right)^{n-2}$$

und daraus durch Integration von 0 bis u wegen $J(0) = 0$

$$(13.30) \quad J(u) = W_{n-1} \left(\frac{\sin \sqrt{K} u}{\sqrt{K}} \right)^{n-1}.$$

³¹⁾ l. c. S. 723.

³²⁾ Vgl. z. B. Schmidt II, III, S. 219 bzw. 770.

Mithin ist mit Rücksicht auf (13. 20), wenn man v an Stelle von ϑ_n schreibt,

$$(13. 31) \quad V = W_{n-1} \int \left(\frac{\sin \sqrt{K} u}{\sqrt{K}} \right)^{n-1} dv.$$

Man kann hier auch s als Parameter einführen und hat dann

$$(13. 32) \quad V = W_{n-1} \int_C \left(\frac{\sin \sqrt{K} u}{\sqrt{K}} \right)^{n-1} \frac{dv}{ds} ds.$$

In allen diesen Formeln bedeutet W_{n-1} das Volumen der $(n-1)$ -dimensionalen Einheitskugel.

Gleichwertige Formeln für O und V wurden unter Zugrundelegung einer anderen Metrik von Erhard Schmidt²³⁾ bewiesen. Interessant ist an den beiden Formeln (13. 17) und (13. 32) zu bemerken, daß sie für $K \rightarrow 0$ in die bekannten Formeln des gewöhnlichen euklidischen Raumes

$$(13. 33) \quad O = (n-1) W_{n-1} \int_C u^{n-2} ds$$

bzw.

$$(13. 34) \quad V = W_{n-1} \int_C u^{n-1} \frac{dv}{ds} ds$$

übergehen.

14. Anwendung auf das isoperimetrische Problem im Raum. Wir wenden uns jetzt zur Lösung der in Nr. 12 gestellten isoperimetrischen Aufgabe:

Es soll unter allen Rotationskörpern R , welche im Fundamentalbereich des Raumes Θ_n konstanter Krümmung mit der Metrik

$$(14. 1) \quad ds^2 = d\vartheta_1^2 + \cos^2 \sqrt{K} \vartheta_1 d\vartheta_2^2 + \dots + \cos^2 \sqrt{K} \vartheta_1 \dots \cos^2 \sqrt{K} \vartheta_{n-1} d\vartheta_n^2$$

liegen, im Zylinder \mathfrak{S}_s ²⁴⁾

$$(14. 2) \quad \frac{\cos^2 \sqrt{K} \vartheta_1 \cos^2 \sqrt{K} \vartheta_2 \dots \cos^2 \sqrt{K} \vartheta_{n-1}}{K} \geq \frac{\cos^2 \sqrt{K} a}{K}$$

einbeschrieben sind und eine vorgegebene Oberfläche O besitzen, derjenige Körper bestimmt werden, welcher das Maximum des eingeschlossenen Inhalts aufweist.

Damit die Aufgabe überhaupt lösbar ist, muß O die Ungleichung

$$(14. 3) \quad O > 2(n-1) W_{n-1} \int_0^a \left(\frac{\sin \sqrt{K} u}{\sqrt{K}} \right)^{n-2} du = O^*$$

erfüllen und im Falle $K > 0$ noch einigen Ungleichungen nach oben genügen, falls eine einfache Bedeckung von G_0 durch die Extremalkörper gefordert wird. Wir werden darauf nochmals zurückkommen.

²³⁾ Vgl. z. B. Schmidt II, III, S. 219 bzw. 774.

²⁴⁾ Für $K = 0$ im Zylinder (12. 5).

Mit Hilfe von (13. 17) und (13. 32) kann diese Aufgabe auf die in Nr. 12 schon gelöste Aufgabe zurückgeführt werden, und zwar mit

$$(14. 4) \quad G = \int_{c_1} \left(\frac{\sin \sqrt{K} u}{\sqrt{K}} \right)^{n-2} ds$$

und

$$(14. 5) \quad J = \int_{c_1} \left(\frac{\sin \sqrt{K} u}{\sqrt{K}} \right)^{n-1} \frac{dv}{ds} ds.$$

Man braucht bloß die Rotationskurve C von Nr. 12 durch Spiegelung an der v -Linie ($u = 0$) zu ergänzen, um dadurch eine einfache, geschlossene, in

$$(14. 6) \quad -a \leq u \leq a$$

einbeschriebene und symmetrische Kurve \bar{C} zu erhalten. Nach den Ausführungen von Nr. 12 fallen die Extremalen \bar{C} unseres Problems mit derjenigen der klassischen isoperimetrischen Aufgabe zusammen, und somit ist das räumliche Problem auf dasjenige der Ebene zurückgeführt worden. Die hier auftretenden Extremalgebiete können dann durch Rotation der Extremalen der klassischen, ebenen, isoperimetrischen Aufgabe erhalten werden. Die allgemeine isoperimetrische Gleichung für den Raum kann man aus (10. 13) erhalten, wenn man G, J durch (14. 4), (14. 5) ersetzt.

Nun ist

$$(14. 7) \quad W_{n-1}J = V, \quad W_{n-1}J_\gamma = V_\gamma$$

und

$$(14. 8) \quad (n-1)W_{n-1}G = O, \quad (n-1)W_{n-1}G_\gamma = O_\gamma,$$

es wird also, wenn man noch

$$(14. 9) \quad (n-1)W_{n-1}R_n(C, C_\gamma) = R_n(C, C_\gamma)$$

setzt,

$$(14. 10) \quad (n-1)(V - V_\gamma) = \varrho_\gamma(O - O_\gamma) - 2 \varrho_\gamma R_n(C, C_\gamma).$$

Das Restglied $R_n(C, C_\gamma)$ kann man auch in der Form

$$(14. 11) \quad R_n(C, C_\gamma) = \int \sin^2 \frac{(\mathfrak{n}, \tilde{\mathfrak{n}})}{2} dO$$

schreiben. Die Zuordnung der Normalen \mathfrak{n} von \mathfrak{R} und der Normalen $\tilde{\mathfrak{n}}$ von \mathfrak{R}_γ geschieht hier nach dem für $n = 2$ geschilderten Vorgang.

Aus der (10. 13) entsprechenden isoperimetrischen Gleichung für „Kreis-Rechteck-Extremalen“ erhält man ebenfalls

$$(14. 12) \quad (n-1)(V - V_q^*) = \varrho_0(O - O_q^*) - 2 \varrho_0 R_n(\mathfrak{R}, \mathfrak{R}_q^*).$$

Dabei wird der Extremalkörper vom „Kugel-Zylinder Typus“ durch Rotation einer „Kreis-Rechteck-Extremale“ O_q^* erzeugt. Für $q = 0$ fallen beide Gleichungen (14. 10), (14. 12) mit der Gleichung

$$(14. 13) \quad (n-1)(V - V_0) = \varrho_0(O - O_0) - 2 \varrho_0 R_n(R, R_0)$$

zusammen.

In meiner vorhin zitierten Arbeit habe ich diese isoperimetrischen Gleichungen, welche in Form von Ungleichungen (ohne die Restglieder) zuerst von Erhard Schmidt bewiesen wurden, auf anderem Wege abgeleitet und namentlich über den Gaußschen Divergenzatz. Der Weg war dort zwar komplizierter als der hier beschrittene, hatte jedoch den Vorteil, eine größere Klasse zulässiger Körper zu umfassen. Da aber für das allgemeine isoperimetrische Problem in Riemannschen Räumen konstanter Krümmung die Einschränkung auf Rotationskörper genügt, so ist die hier gemachte Einschränkung für das allgemeine Problem nicht nachteilig. Die Zurückführung des allgemeinen Problems auf ein solches, in dem nur Rotationskörper zugelassen zu werden brauchen, kann, wie dies von Erhard Schmidt gezeigt wurde, mit Hilfe einer Erweiterung der Schwarzschen Abrundung durchgeführt werden.

Um nun den allgemeinen Satz aussprechen zu können, bedürfen wir des Nachweises, daß O_γ bzw. O_q^* monotone Funktionen von γ bzw. q sind. Für O_q^* ist dies trivial und folgt aus der Formel³⁵⁾

$$(14. 14) \quad O_q^* = O_0 + 2(n-1)W_{n-1}q \left(\frac{\sin \sqrt{K}a}{\sqrt{K}} \right)^{n-2} \cos \sqrt{K}a.$$

Für O_γ folgt dies aus (10. 16) unmittelbar. Man findet

$$(14. 15) \quad O_\gamma = 2(n-1)W_{n-1} \int_0^a \left(\frac{\sin \sqrt{K}u}{\sqrt{K}} \right)^{n-2} \frac{du}{\sqrt{1 - k^2 \frac{\operatorname{tg}^2 \sqrt{K}u}{K}}},$$

Somit ist bewiesen, daß für $K \leq 0$ zu einem vorgegebenen O , das (14. 3) erfüllt, eine und nur eine Extremalfläche konstruiert werden kann. Das Maximum des eingeschlossenen Volumens (da das Restglied nur dann verschwindet, wenn R bis auf Parallelverschiebung mit einem R_γ bzw. einem R_q^* übereinstimmt) wird dann, wie aus (14. 10), (14. 12) folgt, nur für diese Extremalfläche erreicht. Die Art der Extremalfläche wird nur durch O bestimmt.

Für $K > 0$ gelten dieselben Aussagen wie für $K \leq 0$ mit dem einzigen Unterschied, da der Fundamentalbereich beschränkt ist, daß O nach oben

³⁵⁾ Dinghas, I. c. S. 722.

beschränkt ist. Sollen hier nur Körper zugelassen werden, die in G_0 liegen, so muß für den elliptischen Fall

$$(14.16) \quad O < O_0 + 2(n-1) W_{n-1} \left(\frac{\pi}{2\sqrt{K}} - a \right) \left(\frac{\sin \sqrt{K} a}{\sqrt{K}} \right)^{n-2} \cos \sqrt{K} a = O_1$$

und für den sphärischen Fall

$$(14.17) \quad O < O_0 + 2(n-1) W_{n-1} \left(\frac{\pi}{\sqrt{K}} - 2a \right) \left(\frac{\sin \sqrt{K} a}{\sqrt{K}} \right)^{n-2} \cos \sqrt{K} a = O_2$$

bleiben. Für

$$(14.18) \quad O_0 \leq O < O_1 \quad (\text{bzw. } < O_2)$$

ist der Extremalkörper ein solcher vom Kugel-Zylinder-Typus. Für

$$(14.19) \quad O^* < O < O_0$$

ist dieser vom linsenförmigen Typus. Beide Typen werden bestimmt durch die Gleichung $O_\gamma = O$ bzw. $O_\gamma^* = O$.

Wir fassen die Resultate dieser Nummer im folgenden Satz zusammen, deren Inhalt auf Erhard Schmidt zurückgeht:

Die zuletzt gestellte Aufgabe hat für

$$(14.20) \quad O > 2(n-1) W_{n-1} \int_0^a \left(\frac{\sin \sqrt{K} u}{\sqrt{K}} \right)^{n-2} du = O^*$$

eine und nur eine Lösung. Die gesuchten Extremalen sind vom linsenförmigen bzw. Kugel-Zylinder-Typus, je nachdem O die Ungleichung

$$(14.21) \quad O^* < O < O_0$$

oder

$$(14.22) \quad O_0 \leq O$$

erfüllt. Für $K > 0$ muß O außerdem je nach dem Fall (14.16) oder (14.17) genügen.

Mit der Beantwortung der vorhin gestellten Aufgabe ist zugleich auch die in der Einleitung gestellte und mit dieser äquivalente Aufgabe unter Zugrundelegung der Cayleyschen Metrik gelöst. Wir verzichten hier auf die ausführliche Wiedergabe des Wortlautes der entsprechenden Sätze, da ja alles aus dem Fall $n = 2$ entnommen werden kann. Die gesuchten Extremalflächen werden durch Rotation der linsenförmigen bzw. der Kreis-Rechteck-Extremalen um ihre Symmetrieachse X_n erzeugt.

Was die Verallgemeinerung der Resultate von Nr. 9 auf den Raum anbetrifft, so kann hier nichts Näheres gesagt werden. Eine isoperimetrische Ungleichung im Raum von der Form (9.19) ist zur Zeit nicht bekannt und es

scheint, daß diese, falls sie existiert, für ungerades n eine transzendente und für gerades n eine algebraische Form besitzen muß. Jedenfalls wäre die Bedeutung solcher Verallgemeinerungsversuche für die vorliegende Arbeit nicht groß gewesen, da hier nur von Rotationskörpern die Rede gewesen ist. Es ist einleuchtend, daß solche Übertragungen nur im Zusammenhang mit einem Symmetrisierungsverfahren ein relatives Interesse beanspruchen können.

Als letzte Bemerkung möchte ich noch hinzufügen, daß die in meiner vorhin erwähnten Arbeit bewiesene Ungleichung

$$(14.23) \quad V \leq n W_n \int_0^1 \frac{x^{n-1} dx}{\sqrt{1-Kx^2}}$$

(welche dem Inhalt nach auf Erhard Schmidt zurückgeht) für den sphärischen, hyperbolischen und euklidischen Raum, sowie auch für den elliptischen Raum gilt, sobald die Ungleichung $O < n W_n$ erfüllt ist. Das ist insbesondere für jeden Körper der Fall, dessen Punkte alle im Endlichen liegen.

(Eingegangen am 17. 4. 1942.)

Über die Flächenmaße im Euklidischen Raum.

Von

Georg Nöbeling in Erlangen.

In einer kürzlich erschienenen Arbeit¹⁾ haben wir gezeigt, daß die zahlreichen bekannten eindimensionalen Maße von Mengen des Euklidischen Raumes E_n , von einer einzigen Ausnahme abgesehen, für jedes Kontinuum K des E_n gleich einer und derselben Zahl $L(K) \leq +\infty$ sind, die wir die Länge von K genannt haben. Ein entsprechender Satz für die zweidimensionalen Maße ist auch in bescheidenstem Umfange nicht richtig. So ist bereits bekannt, daß 1. die nach unten halbstetigen Maße für Flächen im Raum von Lebesgue, Radó und Menger²⁾, die in gewisser Hinsicht den Charakter innerer Maße tragen, sogar für ebene Flächen kleinere Werte liefern können als die äußeren zweidimensionalen Punktmengenmaße im Raum von Lebesgue²⁾, Groß³⁾, Carathéodory³⁾ usw.⁴⁾ und daß 2. die letzteren für nulldimensionale kompakte Mengen im allgemeinen auch untereinander verschieden sind. In dieser Arbeit wollen wir zeigen, daß die bekannteren äußeren zweidimensionalen Maße sogar für Flächen im allgemeinen auch untereinander verschieden sind.

¹⁾ Jahresbericht der DMV 52 (1942), S. 132—160.

²⁾ H. Lebesgue, *Intégrale, longueur, aire*. *Annali di Matematica* (3) 7 (1902), S. 231—359; T. Radó, *Math. Annalen* 100 (1928), S. 445—479; K. Menger, *Ergebnisse eines math. Kolloquiums*, Heft 2 (1932), S. 10.

³⁾ W. Groß, *Monatsh. f. Math. u. Phys.* 29 (1918), S. 145—193; C. Carathéodory, *Göttinger Nachrichten* 1914, S. 404.

⁴⁾ Es sei F eine in einer Ebene des Raumes E_3 liegende Fläche (topologisches Bild einer abgeschlossenen Kreisscheibe), für welche die Randkurve R einen positiven ebenen Lebesgueschen Inhalt $m^{(2)}R$ hat; es sei \bar{F} das Innere von F . Dann ist das Lebesguesche Flächenmaß $\mu_L(F) = m^{(2)}\bar{F}$, also $< m^{(2)}F$ [Groß³⁾]. Man kann übrigens F erweitern zu einer Fläche $G \subset E_3$, so daß $F \subset G$ gilt (also auch R zum Rand von G fremd ist) und $G - F$ Summe abzählbar vieler Dreiecke D_i ist; man konstruiert G etwa so, daß $G - F$ sich schlicht in \bar{F} projiziert und R eine „Rückkehrkante“ ist, also die nahe bei R liegenden Dreiecke D_i sehr flach liegen. Ist m die Inhaltssumme der Dreiecke D_i , so ist $\mu_L(G) = m + m^{(2)}\bar{F}$. Die Rückkehrkante R der Fläche G wird also durch μ_L nicht mitgemessen.

§ 1.

Ein allgemeiner Satz über zweidimensionale Maßfunktionen.

Wir beweisen zunächst folgenden Satz.

Satz 1. *Jede kompakte nulldimensionale⁵⁾ Menge P des E_3 ist enthalten in einer Fläche F des E_3 derart, daß $F - P$ Summe abzählbar vieler Dreiecke ist⁶⁾, deren elementargeometrische Flächeninhalte eine endliche Summe haben.*

Hierbei kann man unter einer Fläche nach Belieben das topologische Bild der offenen oder abgeschlossenen Kreisscheibe oder der Kugelfläche verstehen.

Beweis. Wir setzen $i_1 = 1$ und bezeichnen mit W_{i_1} einen offenen, P enthaltenden Würfel des E_3 . Dann ist für $k = 1$ die folgende Induktionsvoraussetzung erfüllt:

a) P ist überdeckt durch endlich viele offene und zusammenhängende Mengen $W_{i_1 \dots i_k}$ mit Durchmessern $< \frac{1}{k-1}$; die abgeschlossenen Hüllen $\bar{W}_{i_1 \dots i_k}$ sind paarweise fremd und die Begrenzungen $B(W_{i_1 \dots i_k})$ sind Summen endlich vieler Dreiecke.

Nach dem Borelschen Überdeckungssatz können wir P überdecken durch endlich viele offene Mengen U_1, U_2, \dots, U_l , deren Begrenzungen $B(U_i)$ zu P fremd, deren abgeschlossene Hüllen in $\sum_{i_1 \dots i_k} W_{i_1 \dots i_k}$ enthalten und deren

Durchmesser $d(U_i) < \frac{1}{k}$ sind. Dieselben Eigenschaften haben dann auch die Mengen $V_1 = U_1, V_2 = U_2 - \bar{U}_1, \dots, V_l = U_l - (\bar{U}_1 + \dots + \bar{U}_{l-1})$, die außerdem paarweise fremd sind. Es seien V'_1, V'_2, \dots, V'_h die (nach dem Borelschen Überdeckungssatz endlich vielen) zu P nicht fremden Komponenten der Mengen V_1, V_2, \dots, V_l . Diese Mengen V'_i haben die genannten Eigenschaften der Mengen U_i und V_i , sind aber außerdem zusammenhängend. Schließlich sei V''_i eine zusammenhängende, offene Summe endlich vieler Würfel, welche den Durchschnitt $P \cdot V'_i$ überdeckt und deren abgeschlossene Hülle in V'_i enthalten ist: eine solche Summe existiert auf Grund des Borelschen Überdeckungssatzes und der Tatsache, daß V'_i zusammenhängend und der Durchschnitt $P \cdot V'_i$ wegen $P \cdot B(V'_i) = 0$ kompakt ist. Wir bezeichnen nun die in $W_{i_1 \dots i_k}$ enthaltenen unter den Mengen V''_i in irgendeiner Reihenfolge mit $W_{i_1 \dots i_k i_{k+1}}$ ($i_{k+1} = 1, 2, \dots$). Dann erfüllen die Mengen $W_{i_1 \dots i_k i_{k+1}}$,

⁵⁾ Eine Menge P des E_3 heißt kompakt, wenn sie beschränkt und abgeschlossen ist. Sie heißt nulldimensional, wenn jeder Punkt von P enthalten ist in beliebig kleinen Umgebungen U , deren Begrenzungen $B(U)$ zu P fremd sind (vgl. K. Menger, Dimensionstheorie, Leipzig 1928).

⁶⁾ Und zwar bilden diese Dreiecke eine Triangulierung von $F - P$ im Sinne der Topologie.

da sie mit den Mengen V_i'' identisch sind, die Bedingungen a) mit $k+1$ statt k , und außerdem gilt:

$$b) \quad \bar{W}_{i_1 \dots i_k i_{k+1}} \subset W_{i_1 \dots i_k}.$$

Wir können also a) und b) für jedes natürliche k als erfüllt annehmen.

Nun sei weiter T eine zum P enthaltenden, offenen Würfel W_{i_1} fremde Tetraederfläche, deren Durchschnitt mit der Begrenzung $B(W_{i_1})$ von W_{i_1} ein Dreieck \bar{D}_{i_1} mit einem Flächeninhalt $< \frac{1}{2^3}$ ist; das Innere dieses Dreiecks bezeichnen wir mit D_{i_1} oder mit E_{i_1} . Zu jeder Menge $W_{i_1 \dots i_k}$ wählen wir in D_{i_1} ein offenes Dreieck $D_{i_1 \dots i_k}$ mit einem Durchmesser $< \frac{1}{k-1}$ derart, daß je zwei dieser Dreiecke mit gleich vielen Indizes einen positiven Abstand haben und so, daß außerdem $\bar{D}_{i_1 \dots i_k i_{k+1}} \subset D_{i_1 \dots i_k}$ ist. — Schließlich sei Q die Menge aller Punkte von T , die in unendlich vielen Dreiecken $D_{i_1 \dots i_k}$ liegen; also $Q = \prod_k \sum_{i_1 \dots i_k} D_{i_1 \dots i_k}$.

Wir beginnen jetzt die Konstruktion der gesuchten, P enthaltenden Fläche F , und zwar konstruieren wir F als topologisches Bild $t(T)$ der Tetraederfläche T . Diese Abbildung t wird Schritt für Schritt auf den Mengen Q , $T - D^1, D^1 - D^2, D^2 - D^3, \dots$ konstruiert, die paarweise fremd sind und deren Summe gleich T ist; dabei bedeutet D^k die Summe $\sum_{i_1 \dots i_k} D_{i_1 \dots i_k}$ aller Dreiecke mit k Indizes.

Jeder Punkt q von Q kann auf genau eine Art als Durchschnitt von Dreiecken $D_{i_1 \dots i_k}$ dargestellt werden: $(q) = D_{i_1} \cdot D_{i_1 i_2} \cdot D_{i_1 i_2 i_3} \dots$. Diesem Punkt q ordnen wir den Punkt p der nulldimensionalen Menge P mit $(p) = W_{i_1} \cdot W_{i_1 i_2} \cdot W_{i_1 i_2 i_3} \dots$ als Bild $t(q)$ zu. Aus den Eigenschaften a) und b) der $W_{i_1 \dots i_k}$ und den entsprechenden Eigenschaften der Dreiecke $D_{i_1 \dots i_k}$ folgt unmittelbar, daß die Abbildung t von Q auf P topologisch ist.

Auf $T - D^1$ setzen wir t gleich der identischen Abbildung.

Zu jeder der endlich vielen Mengen $W_{i_1 i_2}$ mit zwei Indizes wählen wir eine Menge $O_{i_1 i_2}$ mit folgenden Eigenschaften: 1. $O_{i_1 i_2}$ ist homöomorph zum Zylinder $x^2 + y^2 = 1, 0 \leq z \leq 1$ und Summe endlich vieler Dreiecke; 2. $O_{i_1 i_2}$ ist in $W_{i_1} - \sum_{i_3} \bar{W}_{i_1 i_2 i_3}$ enthalten; 3. die eine Randkurve $R'_{i_1 i_2}$ von $O_{i_1 i_2}$ ist die Begrenzung eines offenen Dreiecks $E'_{i_1 i_2} \subset E_{i_1}$, die andere Randkurve $R_{i_1 i_2}$ ist die Begrenzung eines offenen Dreiecks $E_{i_1 i_2} \subset B(W_{i_1 i_2})$; 4. je zwei Mengen $O_{i_1 i_2}$ sind fremd; 5. die Gesamtoberfläche aller $O_{i_1 i_2}$ ist $< \frac{1}{2^2}$, die aller $E_{i_1 i_2}$ ist $< \frac{1}{2^3}$. — Diese Mengen $O_{i_1 i_2}$ sind also sehr dünne,

⁷⁾ Allgemein bezeichnen wir mit \bar{M} die Menge der inneren Punkte von M ; in diesem Fall also die nicht auf den Randkurven liegenden Punkte von $O_{i_1 i_2}$.

paarweise fremde Schläuche, welche in $\bar{W}_{i_1} - \sum_{i_2} W_{i_1 i_2}$ verlaufen und das Dreieck D_{i_1} mit den Oberflächen $B(W_{i_1 i_2})$ von $W_{i_1 i_2}$ verbinden. — Wir bilden nun $D^1 - D^2$ topologisch auf die Menge $(E_{i_1} - \sum_{i_2} E_{i_1 i_2}) + \sum_{i_2} O_{i_1 i_2}$ ab derart, daß diese Abbildung t sich stetig anschließt an die bereits auf $T - D^1$ definierte Abbildung t und so daß außerdem der Rand von $D_{i_1 i_2}$ auf den Rand $R_{i_1 i_2}$ von $O_{i_1 i_2}$ abgebildet wird. Das Bild $t(T - D^1)$ ist Summe endlich vieler Dreiecke und hat einen elementargeometrischen Flächeninhalt $< \frac{1}{2}$, da E_{i_1} und $\sum_{i_2} O_{i_1 i_2}$ Flächeninhalte $< \frac{1}{2^2}$ haben.

Analog definiert man t auf $D^2 - D^3$, indem man aus jedem Dreieck $E_{i_1 i_2}$ kleine dreieckige Löcher $E'_{i_1 i_2 i_3}$ herausschneidet und an jedes dieser Löcher einen dünnen Schlauch $O_{i_1 i_2 i_3}$ ansetzt, dessen zweites Ende der Rand eines Dreiecks $E_{i_1 i_2 i_3}$ in der Begrenzung $B(W_{i_1 i_2 i_3})$ ist; dabei sollen diese Schläuche paarweise fremd sein, es soll $O_{i_1 i_2 i_3}$ in $\bar{W}_{i_1 i_2} - \sum_{i_3} \bar{W}_{i_1 i_2 i_3}$ enthalten sein und die Flächeninhalte aller $O_{i_1 i_2 i_3}$ sollen zusammen $< \frac{1}{2^3}$, die aller Dreiecke $E'_{i_1 i_2 i_3}$ zusammen $< \frac{1}{2^4}$ sein. Nun wird t auf $D^2 - D^3$ stetig fortgesetzt, indem $D^2 - D^3$ topologisch auf die Menge $(\sum_{i_2} E_{i_1 i_2} - \sum_{i_2 i_3} E_{i_1 i_2 i_3}) + \sum_{i_2 i_3} O_{i_1 i_2 i_3}$ abgebildet wird, die einen Flächeninhalt $< \frac{1}{2^3}$ hat, da die Mengen $\sum_{i_2} E_{i_1 i_2}$ und $\sum_{i_2 i_3} O_{i_1 i_2 i_3}$ Flächeninhalte $< \frac{1}{2^3}$ haben.

Durch Fortsetzung dieses Verfahrens wird t auf jeder Menge $D^k - D^{k+1}$ und damit auf ganz T definiert. Der stetige Anschluß an die bereits auf Q definierte Abbildung t wird dadurch gewährleistet, daß jeder Schlauch $O_{i_1 \dots i_k}$ in $\bar{W}_{i_1 \dots i_k}$ enthalten ist. Die Eineindeutigkeit von t ergibt sich daraus, daß t auf Q , $T - D^1$ und jedem $D^k - D^{k+1}$ topologisch ist, die Bilder dieser Mengen in P , $E_3 - W^1$ und $W^k - W^{k+1}$ enthalten und diese Mengen paarweise fremd sind (W^k ist dabei die Summe aller Mengen $W_{i_1 \dots i_k}$ mit k Indizes). Also ist $t(T)$ eine P enthaltende, einfach zusammenhängende Fläche F . Die Menge $F - P$ ist Summe von abzählbar vielen Dreiecken wegen $F - P = (T + \sum O_{i_1 \dots i_k} + \sum E_{i_1 \dots i_k}) - \sum E'_{i_1 \dots i_k i_{k+1}}$. Die Inhaltssumme dieser Dreiecke ist endlich, da der Flächeninhalt von $t(D^k - D^{k+1}) < \frac{1}{2^k}$ ist und auch $T - D^1$ einen endlichen Flächeninhalt hat.

F erfüllt also alle Behauptungen des Satzes 1. Läßt man aus F ein zu P fremdes abgeschlossenes oder offenes Dreieck weg, so erhält man eine P enthaltende Fläche, die zur offenen bzw. abgeschlossenen Kreisscheibe homöomorph ist und ebenfalls den Behauptungen des Satzes 1 genügt. Satz 1 ist also vollständig bewiesen.

Satz 2. (Hauptsatz). *Es seien μ_1 und μ_2 zweidimensionale Maßfunktionen⁹⁾ im E_3 . Es existiere im E_3 eine kompakte, nulldimensionale Menge P , für welche $\mu_1(P) < \mu_2(P)$ ist. Dann existiert im E_3 eine Fläche F , für welche $\mu_1(F) < \mu_2(F)$ ist [dabei ist auch $\mu_2(F) < \infty$, wenn $\mu_2(P) < \infty$ ist].*

Beweis. Wir legen durch P eine Fläche F im Sinne des Satzes 1. Da $F - P$ eine Summe von Dreiecken mit einer endlichen Inhaltssumme ist und für jedes dieser Dreiecke die Maße μ_1 und μ_2 mit seinem elementargeometrischen Inhalt übereinstimmen, ist $\mu_1(F - P) = \mu_2(F - P) < \infty$. Hieraus und aus $\mu_1(P) < \mu_2(P)$ folgt die Behauptung.

§ 2.

Die betrachteten äußeren zweidimensionalen Maße.

A. Projektionsmaße.

I. O. Janzen⁹⁾ definiert ein zweidimensionales Maß folgendermaßen. Es seien feste rechtwinklige Koordinaten x_1, x_2, x_3 im E_3 gegeben. Für jedes natürliche k ist der E_3 Summe der paarweise fremden, halboffenen Würfel $(z_i - 1)/2^k \leq x_i < z_i/2^k$ ($i = 1, 2, 3$), wobei die z_1, z_2, z_3 unabhängig voneinander alle Zahlen $0, \pm 1, \pm 2, \dots$ durchlaufen. Diese Würfel bezeichnen wir in irgendeiner Reihenfolge mit W_k^1, W_k^2, \dots . Es sei nun M eine Menge des E_3 . Dann bezeichnen wir mit M_{kj}^i die Projektion des Durchschnitts $M_k^i = W_k^i \cdot M$ auf die drei Koordinatenebenen $x_j = 0$ ($j = 1, 2, 3$), mit $m^{(2)} M_{kj}^i$ das äußere Lebesguesche Maß von M_{kj}^i und setzen

$$T_k(M) = \sum_{i=1}^{\infty} \sqrt{(m^{(2)} M_{k1}^i)^2 + (m^{(2)} M_{k2}^i)^2 + (m^{(2)} M_{k3}^i)^2}.$$

Da die Würfel W_{k+1}^i aus den Würfeln W_k^i durch Unterteilung hervorgehen, gilt $T_k(M) \leq T_{k+1}(M)$. Mithin existiert $\lim T_k(M) = \mu_j(M) \leq \infty$.

II. W. Groß¹⁰⁾ projiziert die Menge M_k^i auf eine beliebige Ebene E des E_3 und bestimmt das äußere Lebesguesche Maß dieser Projektion. Es sei

⁹⁾ Eine zweidimensionale Maßfunktion $\mu(M)$ des E_3 ist durch folgende fünf Eigenschaften gekennzeichnet (vgl. C. Carathéodory, Vorlesungen über reelle Funktionen Leipzig 1918, S. 238): I. Die Zahl $\mu(M)$, die jeder beliebigen Punktmenge M des E_3 zugeordnet wird, ist entweder Null, oder endlich und positiv, oder gleich $+\infty$. Es gibt Punktmenge, für welche diese Zahl > 0 und endlich ist; für die leere Menge ist sie gleich Null. II. Für eine Teilmenge N von M ist stets $\mu(N) \leq \mu(M)$. III. Ist V die Vereinigungsmenge einer Folge von endlich oder abzählbar unendlich vielen Punktmenge M_1, M_2, \dots , so ist stets $\mu(V) \leq \mu(M_1) + \mu(M_2) + \dots$. IV. Sind M und N zwei Punktmenge, deren Entfernung $\delta > 0$ ist, so ist stets $\mu(M + N) = \mu(M) + \mu(N)$. V. Für eine Dreiecksfläche D ist $\mu(D)$ gleich dem elementargeometrischen Inhalt.

¹⁰⁾ „Über einige stetige Kurven, über Bogenlänge, linearen Inhalt und Flächeninhalt“, Inaugural-Dissertation, Königsberg 1907.

¹¹⁾ A. a. O.⁹⁾, S. 156.

$G(M_k^i)$ die obere Grenze dieser Maße für alle Ebenen E . Setzt man $S_k(M) = \sum_{i=1}^{\infty} G(M_k^i)$, so gilt $S_k(M) \leq S_{k+1}(M)$. Also existiert $\lim S_k(M) = \mu_G(M)$. — Übrigens ist $\mu_G(M)$ die kleinste Maßfunktion $\mu(M)$, welche den Axiomen I–IV von C. Carathéodory¹¹⁾ und außerdem der folgenden Forderung genügt: $\mu(M)$ ist nicht kleiner als das äußere Lebesguesche Maß $m^{(2)}M_E$ der Projektion M_E von M in eine beliebige Ebene E .

B. Überdeckungsmaße.

III. C. Carathéodory¹¹⁾ überdeckt die gegebene Menge M des E_3 durch endlich oder abzählbar unendlich viele konvexe Mengen K_1, K_2, \dots mit Durchmessern $< \varrho$ ($\varrho > 0$). Es sei $d^{(2)}(K_i)$ der zweidimensionale Durchmesser von K_i , d. h. die obere Grenze der Lebesgueschen Inhalte der Projektionen von K_i in alle Ebenen E . Ist nun λ_ϱ die untere Grenze der Summen $d^{(2)}(K_1) + d^{(2)}(K_2) + \dots$ für alle Überdeckungen der genannten Art, so existiert $\lim_{\varrho \rightarrow 0} \lambda_\varrho = \mu_C(M)$.

IV. F. Hausdorff¹²⁾ verfährt genau so, läßt aber für die K_i nur Kugeln zu. Das so entstehende Maß heiße $\mu_H(M)$.

C. Ein integralgeometrisches Maß.

V. J. Favard¹³⁾ bezeichnet, auf M. W. Crofton und H. Lebesgue fußend, für eine beliebige Gerade \mathfrak{G} des E_3 mit $m = m(\mathfrak{G}) \leq \infty$ die Anzahl der Punkte des Durchschnitts $M \cdot \mathfrak{G}$ und führt die Zahl $\mu_F(M) = \frac{1}{\pi} \int m \mathfrak{G}$ als zweidimensionales Maß von M ein, wobei \mathfrak{G} die Geradendichte¹⁴⁾ im E_3 ist.

D. Axiomatische Maße.

VI. A. Kolmogoroff¹⁵⁾ nennt eine nicht negative Mengenfunktion $\mu_K(A)$, welche für alle Suslinschen Mengen A definiert ist, ein zweidimensionales Maß, wenn folgende Axiome erfüllt sind: a) ist A in der Summe der endlich oder abzählbar unendlich vielen Mengen A_1, A_2, \dots enthalten, so gilt $\mu_K(A)$

¹¹⁾ Göttinger Nachrichten 1914, S. 404.

¹²⁾ Math. Annalen 79 (1919), S. 163.

¹³⁾ C. R. Acad. Sci., Paris 194 (1932), S. 344.

¹⁴⁾ Vgl. z. B. W. Blaschke, Vorlesungen über Integralgeometrie, 2. Heft, 1937, S. 66.

¹⁵⁾ Math. Annalen 107 (1933), S. 351. Zur Definition der analytischen oder Suslinschen Mengen vgl. z. B. F. Hausdorff, Mengenlehre, 1927. Hier möge die Bemerkung genügen, daß jede Borelsche Menge, insbesondere also jede abgeschlossene Menge des E_3 analytisch ist. Eine Abbildung heit dehnungslos, wenn der Abstand je zweier Punkte der Urbildmenge größer oder gleich dem Abstand der Bildpunkte ist.

$\leq \mu_K(A_1) + \mu_K(A_2) + \dots$; b) sind die endlich oder abzählbar unendlich vielen Teilmengen A_1, A_2, \dots von A paarweise fremd, so gilt $\mu_K(A_1) + \mu_K(A_2) + \dots \leq \mu_K(A)$; c) ist A' ein dehnungsloses Bild von A , so ist $\mu_K(A') \leq \mu_K(A)$; d) für das Einheitsquadrat Q ist $\mu_K(Q) = 1$.

§ 3.

Flächen mit nicht zusammenfallenden Maßen.

Satz 3. *Es gibt im E_3 eine Fläche F , für welche $\mu_J(F) < \mu_G(F)$ ist.*

Beweis. Im Einheitsintervall $0 \leq x_i \leq 1$ der x_i -Achse wählen wir je ein Cantorsches Diskontinuum D_i ($i = 1, 2$), in der x_3 -Achse ein Diskontinuum D_3 mit positivem linearem Lebesgueschem Inhalt $m^{(1)}D_3$. Es sei P die perfekte, nulldimensionale Menge aller Punkte (x_1, x_2, x_3) des E_3 mit $x_i \in D_i$ ($i = 1, 2, 3$). Da das Cantorsche Diskontinuum einen verschwindenden linearen Lebesgueschen Inhalt hat, hat die Projektion von P auf jede der drei Ebenen $x_i = 0$ ($i = 1, 2, 3$) den ebenen Lebesgueschen Inhalt Null. Also ist $\mu_J(P) = 0$. Hingegen ist die Projektion von P auf die Ebene $x_1 = x_2$ identisch mit der Menge aller Punkte (x_1, x_2, x_3) des E_3 , für welche (x_1, x_2) in der Diagonalen $x_1 = x_2$, $0 \leq x_i \leq 1$ ($i = 1, 2$) und x_3 in D_3 liegt; diese Projektion hat also den ebenen Lebesgueschen Inhalt $\sqrt{2} m^{(1)}D_3 > 0$. Mithin ist $\mu_G(P) > 0$. Aus der damit bewiesenen Ungleichung $\mu_J(P) < \mu_G(P)$ folgt jetzt die Behauptung des Satzes 3 mit Hilfe von Satz 2.

Satz 4. *Es gibt im E_3 eine Fläche F , für welche die Ungleichungen $\mu_J(F) < \mu_G(F)$ und $\mu_F(F) < \mu_G(F)$ bestehen.*

Beweis. Wir konstruieren¹⁶⁾ in der x_1x_2 -Ebene E_2 des E_3 zunächst eine kompakte, nulldimensionale Menge Q mit folgenden Eigenschaften: a) die Projektion von Q in die x_1 -Achse hat einen positiven linearen Lebesgueschen Inhalt; b) die Projektion von Q in jede zur x_1 -Achse nicht parallele Gerade $\subset E_2$ ist eine lineare Lebesguesche Nullmenge.

Es sei Q_1 das Einheitsquadrat $0 \leq x_1 \leq 1$, $0 \leq x_2 \leq 1$. Wir nennen ein in Q_1 enthaltenes, abgeschlossenes Parallelogramm regulär, wenn zwei seiner Seiten zur x_2 -Achse parallel sind; nur diese beiden Seiten sollen fortan Seiten heißen; die beiden anderen wollen wir Basen nennen. Eine Summe S endlich vieler regulärer Parallelogramme möge ausgezeichnet heißen, wenn jede Gerade $x_1 = c$ mit $0 \leq c \leq 1$ entweder mit genau einem Parallelogramm von S einen nichtleeren Durchschnitt hat oder von genau zwei Parallelogrammen je eine Seite enthält.

¹⁶⁾ Die Konstruktion ist eine Nachbildung der Konstruktion von S. Saks¹⁷⁾.

Es sei S eine ausgezeichnete Parallelogrammsumme und G eine Gerade der E_2 durch den Ursprung, die mit der x_1 -Achse einen Winkel $w \geq \frac{\pi}{n}$ einschließt (n eine natürliche Zahl ≥ 2). Dann läßt sich eine ausgezeichnete Parallelogrammsumme S' mit folgenden Eigenschaften angeben: c) $S' \subset S$; d) die Basen von S' sind senkrecht zu G ; e) die Seiten von S' haben eine Längensumme $< \frac{1}{n}$. — (Zur Konstruktion von S' verwende man etwa folgende

Bemerkung: In jedem Parallelogramm von S kann man die Seiten verbinden durch einen in diesem Parallelogramm enthaltenen Streckenzug, dessen Strecken zu G oder zur x_1 -Achse senkrecht sind.)

Wir behaupten nun: f) Ist H eine Gerade der E_2 , die mit G einen Winkel $\leq \frac{\pi}{n^2}$ einschließt, so hat die Projektion S'_H von S' in H einen linearen Inhalt $< \frac{5}{n}$. Zum Beweis schätzen wir zunächst die Längensumme l der Basen von S' nach oben ab. Die Basen bilden mit der x_1 -Achse den Winkel $\frac{\pi}{2} - w \leq \frac{\pi}{2} - \frac{\pi}{n}$; ihre Projektionen in die x_1 -Achse überdecken das Einheitsintervall bis auf endlich viele Punkte genau zweifach, da S' ausgezeichnet ist; also ist $l \cdot \sin \frac{\pi}{n} \leq 2$, woraus wegen $\sin \frac{\pi}{n} \geq \frac{\pi}{2n}$ die Ungleichung $l \leq \frac{4n}{\pi}$ folgt. Da nun die Basen mit H einen Winkel $\geq \frac{\pi}{2} - \frac{\pi}{n^2}$ bilden, hat die Projektion der Basen in H einen Inhalt $\leq l \cdot \sin \frac{\pi}{n^2} \leq \frac{4n}{\pi} \cdot \frac{\pi}{n^2} = \frac{4}{n}$. Wegen e) hat die Projektion der Seiten von S' in H einen Inhalt $< \frac{1}{n}$. Also hat S'_H einen Inhalt $< \frac{4}{n} + \frac{1}{n} = \frac{5}{n}$, wie behauptet.

Das Einheitsquadrat Q_1 ist eine ausgezeichnete Parallelogrammsumme. Angenommen, es wäre bereits eine ausgezeichnete Parallelogrammsumme Q_{n-1} definiert. Es sei nun G_n^i die Gerade $x_2 = x_1 \cdot \operatorname{tg} \frac{m}{n^2} \pi$, wobei $m = n + i$ ist mit $i = 0, 1, \dots, (n-2) \cdot n$. Diese Geraden bilden mit der x_1 -Achse Winkel $\geq \frac{\pi}{n}$ und zu je zwei aufeinanderfolgenden den Winkel $\frac{\pi}{n^2}$. Wir wählen nun gemäß c) bis e) zu $S = Q_{n-1}$ und $G = G_n^0$ eine ausgezeichnete Parallelogrammsumme $S' = Q_n^0$; sodann zu $S = Q_n^0$ und $G = G_n^1$ eine ausgezeichnete Parallelogrammsumme $S' = Q_n^1$ usw. Dann gilt, wenn wir noch $Q_n^{(n-2)n} = Q_n$ setzen, $Q_{n-1} \supset Q_n^0 \supset Q_n^1 \supset \dots \supset Q_n$. Wegen f) folgt hieraus: g) Ist H eine Gerade, die mit der x_1 -Achse einen Winkel $\geq \frac{\pi}{n}$ einschließt, so hat die Projektion von Q_n in H einen Inhalt $< \frac{5}{n}$.

Der Durchschnitt $Q = \prod_{n=1}^{\infty} Q_n$ ist die gesuchte kompakte Menge. Da die Parallelogrammsummen Q_n ausgezeichnet sind, ist die Projektion von Q

in die x_1 -Achse das Einheitsintervall und wegen g) ist die Projektion von Q in jede zur x_1 -Achse nicht parallele Gerade der E_2 eine lineare Nullmenge; aus dem letzteren folgt noch, daß die Projektionen auf die Geraden $x_1 = \pm x_2$ nulldimensional sind und hieraus, daß jeder Punkt von Q in beliebig kleinen Rechtecken mit zu Q fremden Begrenzungen enthalten, also Q nulldimensional ist.

Nun sei L ein Diskontinuum der x_3 -Achse mit positivem linearem Inhalt und P die Menge aller Punkte (x_1, x_2, x_3) des E_3 , für welche (x_1, x_2) in Q und x_3 in L liegt. Dann ist P eine kompakte, nulldimensionale Menge. Wir behaupten: h) Die Projektion von P in die Koordinatenebene $x_3 = 0$ hat einen positiven ebenen Lebesgueschen Inhalt; i) die Projektion von P in jede zur x_1 -Achse nicht parallele Ebene E des E_3 ist eine ebene Nullmenge.

h) folgt unmittelbar aus a) und $m^{(1)}L > 0$. i) ergibt sich so: Es sei H die Schnittgerade der Ebene E mit der x_1x_2 -Ebene und P_E die Projektion von P in E . Dann ist die Projektion P_{EH} von P_E in H identisch mit der Projektion Q_H von Q in H , also P_{EH} wegen b) eine lineare und daher P_E eine ebene Nullmenge.

Aus h) folgt: $\mu_J(P) > 0$ und $\mu_G(P) > 0$. Hingegen folgt aus i) die Gleichung $\mu_F(P) = 0$. Denn jeder Geraden \mathfrak{G} des E_3 sei zugeordnet die zu \mathfrak{G} senkrechte Ebene E durch den Ursprung 0. Ist nun $\dot{\varphi}$ das Flächenelement des sphärischen Bildes der Normalen dieser Ebenen E und $\dot{s}_1\dot{s}_2$ die Punktdichte in E , so ist die Geradendichte $\mathfrak{G} = \dot{\varphi}\dot{s}_1\dot{s}_2^{14}$. Da die zur x_1 -Achse parallelen Ebenen eine Nullmenge bilden, folgt hiernach aus i): $\int m\dot{s}_1\dot{s}_2 = 0$ für jede Ebene E durch 0 außer den Ebenen einer Nullmenge. Also ist $\mu_F(P) = \frac{1}{\pi} \int m \mathfrak{G} = \frac{1}{\pi} \int \left(\int m\dot{s}_1\dot{s}_2 \right) \dot{\varphi} = 0$.

Aus $\mu_J(P) > 0$, $\mu_G(P) > 0$ und $\mu_F(P) = 0$ folgt jetzt die Behauptung des Satzes 4 mit Hilfe des Satzes 2.

Anmerkung. Aus a) für den linearen Fall und aus h) für den zweidimensionalen Fall ergibt sich: Das Favardsche Maß einer Menge kann kleiner sein als das (Lebesguesche) Maß der Projektion der Menge in eine Gerade bzw. eine Ebene.

Satz 5. *Es gibt im E_3 eine Fläche F , für welche die (zusammenfallenden) Maße $\mu_J(F)$, $\mu_G(F)$ und $\mu_F(F)$ klein sind als die Maße $\mu_C(F)$, $\mu_H(F)$ und alle Maße $\mu_K(F)$.*

Beweis. S. Saks¹⁷⁾ hat eine im Kreisring $1 \leq x_1^2 + x_2^2 \leq 2$ der x_1x_2 -Ebene enthaltene perfekte Menge Q mit folgenden Eigenschaften angegeben: a) Die Projektion von Q auf jede Gerade ist eine lineare Lebesguesche Null-

¹⁷⁾ Fund. Math. 9 (1927), S. 16–24.

menge; b) jede Halbgerade vom Ursprung aus hat mit Q einen nichtleeren Durchschnitt.

Die Menge Q ist nulldimensional; denn aus a) folgt, daß jeder Punkt von Q enthalten ist in beliebig kleinen, offenen Rechtecken, deren Seiten (etwa zu den Achsen parallel und) zu Q fremd sind.

Es sei nun R die Menge aller Punkte (x_1, x_2, x_3) des E_3 , für welche der Punkt (x_1, x_2) in Q liegt und für welche $0 \leq x_3 \leq 1$ ist. Wir behaupten: c) Die Projektion R_E von R auf eine beliebige Ebene E des E_3 ist eine ebene Lebesguesche Nullmenge. Ist E zur $x_1 x_2$ -Ebene parallel, so ist R_E zu Q kongruent, also nach a) eine ebene Nullmenge. Es sei nun E nicht zur $x_1 x_2$ -Ebene parallel. Dann projizieren sich die zur x_3 -Achse parallelen Strecken von R in (parallele) Strecken von E , die zur Schnittgeraden H von E mit der $x_1 x_2$ -Ebene senkrecht sind. Also ist $R_{EH} = Q_H$. Nach a) ist Q_H eine lineare Nullmenge, also R_E eine ebene Nullmenge. Damit ist c) bewiesen.

Aus c) folgen unmittelbar die beiden ersten der folgenden Gleichungen: d) $\mu_J(R) = 0$, $\mu_G(R) = 0$, $\mu_F(R) = 0$. Die dritte ergibt sich analog wie auf S. 695. Damit ist d) bewiesen.

Nun sei L eine perfekte, nulldimensionale Teilmenge der Strecke $0 \leq x_3 \leq 1$ der x_3 -Achse mit positivem linearem Lebesgueschem Inhalt: $m^{(1)}L > 0$. Mit P bezeichnen wir die perfekte Menge aller Punkte (x_1, x_2, x_3) des E_3 , für welche der Punkt (x_1, x_2) in Q und x_3 in L liegt. Diese Menge P ist nulldimensional; denn wie oben bewiesen, ist jeder Punkt von Q in beliebig kleinen Rechtecken mit zu Q fremden Begrenzungen enthalten; also ist jeder Punkt von P in beliebig kleinen Quadern mit zu P fremden Begrenzungen enthalten.

Wegen $P \subset R$ folgt aus d):

$$(1) \quad \mu_J(P) = \mu_G(P) = \mu_F(P) = 0.$$

Wir behaupten weiter die Ungleichungen

$$(2) \quad \mu_C(P) > 0, \quad \mu_H(P) > 0, \quad \mu_K(P) > 0.$$

Zum Beweis der ersten Ungleichung betrachten wir für eine beliebige offene oder abgeschlossene Strecke S der x_3 -Achse mit der Länge s die Menge R^S aller Punkte (x_1, x_2, x_3) des E_3 , für welche (x_1, x_2) in Q und x_3 in S liegt. Wir behaupten zunächst: $\mu_C(R^S) > 0$. Überdecken die zu Q nicht fremden, konvexen, ebenen Mengen M_1, \dots, M_k mit Durchmessern $d_1, \dots, d_k < \frac{1}{2}$ die Menge Q , so überdecken ihre Zentralprojektionen M'_i von $(0, 0)$ aus auf die Kreislinie $Z = [x_1^2 + x_2^2 = \frac{1}{4}]$ diese Kreislinie vollständig. Da die M_i außerhalb von Z liegen, ist der Durchmesser d_i von M_i nicht kleiner als die Länge des Kreisbogens M'_i . Also ist $d_1 + \dots + d_k$ nicht kleiner als die

Länge von Z , d. h. $d_1 + \dots + d_k \geq \pi$. Dann haben die Orthogonalprojektionen der M_1, \dots, M_k auf die x_1 - oder die x_2 -Achse eine Längensumme $\geq \pi/\sqrt{2}$. Nun sei R^S überdeckt durch endlich viele konvexe, offene Mengen K_1, \dots, K_k mit Durchmessern $< \frac{1}{2}$. Die Durchschnitte $K_i(x_3)$ der Mengen K_i mit der zur x_1x_2 -Ebene parallelen Ebene E durch den Punkt x_3 der Strecke S sind konvexe ebene Mengen, welche den zu Q kongruenten Durchschnitt $R^S \cdot E$ überdecken. Also haben die Projektionen der $K_i(x_3)$ in die x_1 - oder die x_2 -Achse eine Durchmessersumme $\geq \pi/\sqrt{2}$. Dies gilt für jedes x_3 aus S . Also haben die Projektionen der K_i in die x_1x_3 - oder die x_2x_3 -Ebene eine Inhaltssumme $\geq s\pi/2\sqrt{2}$. Um so mehr haben die zweidimensionalen Durchmesser der K_i eine Summe $\geq s\pi/2\sqrt{2}$. Daher ist auch $\mu_G(R^S) \geq s\pi/2\sqrt{2}$, also $\mu_G(R^S) > 0$, wie behauptet. Natürlich ist $\mu_G(R^S)$ proportional zu s , etwa $\mu_G(R^S) = a \cdot s$ mit $a \neq 0$. Da die Menge L der x_3 -Achse aus dem Einheitsintervall der x_3 -Achse hervorgeht durch Tilgung abzählbar vieler offener Intervalle S_1, S_2, \dots , deren Längensumme $s_1 + s_2 + \dots < 1$ ist, so ist $P = R - (R^{S_1} + R^{S_2} + \dots)$. Also ist $\mu_G(P) = a \cdot (1 - (s_1 + s_2 + \dots)) > 0$.

Aus der hiermit bewiesenen Ungleichung $\mu_G(P) > 0$ folgt die Ungleichung $\mu_H(P) > 0$ auf Grund der folgenden Bemerkung: Es ist $\mu_G(M) \leq \mu_H(M)$ für jede Menge M , da bei der Definition von μ_H weniger Mengen zur Überdeckung von M zugelassen sind als bei der Definition von μ_G .

Die letzte Ungleichung $\mu_K(P) > 0$ braucht nur für das kleinste unter den Maßen μ_K bewiesen zu werden. Dieses Minimalmaß $\underline{\mu}_K$ ist folgendermaßen definiert: Es ist $\underline{\mu}_K(A) = \sup \sum m^{(2)} \pi_i(A_i)$, wobei sich das Supremumzeichen auf alle möglichen endlichen oder unendlichen Folgen von analytischen, paarweise fremden Teilmengen A_i von A bezieht und die π_i dehnungslose Abbildungen in eine Ebene bedeuten. Spezialisiert man in dieser Definition die A_i auf die Durchschnitte von A mit den Würfeln W_k^i und die π_i auf senkrechte Projektionen, so erhält man die Definition von μ_G , in welcher der Limes der monoton wachsenden Folge $S_k(M)$ mit ihrem Supremum identisch ist. Also gilt: e) $\mu_G(A) \leq \underline{\mu}_K(A)$ für jede analytische Menge A . — Es sei P' die Menge aller Punkte von P mit $|x_2| \leq |x_1|$ und h die projektive Abbildung $x'_1 = \frac{1}{x_1}$, $x'_2 = \frac{x_2}{x_1}$, $x'_3 = \frac{x_3}{x_1}$ des E_3 , welche die Ebene $x_1 = 0$ ins Unendliche wirft. Sie transformiert die Menge $1 \leq x_1^2 + x_2^2 \leq 2$, $|x_2| \leq |x_1|$ und damit die Menge P' in eine beschränkte Menge; es sei $h(P') = P''$ gesetzt. Jede der Ebenen $ax_1 = bx_2$ mit $|a| \leq |b|$ hat mit P' einen Durchschnitt, welcher eine aus L durch eine Parallelverschiebung hervorgehende Menge enthält. Diese Ebenen gehen durch h über in die zwischen den zwei festen parallelen Ebenen $x_2 = \pm 1$ gelegenen Ebenen. Jede dieser letzteren Ebenen hat mit P'' einen Durchschnitt, welcher eine aus L durch eine Parallelverschiebung

hervorgehende Menge enthält. Nun war L als eine Menge mit positivem linearem Inhalt gewählt worden. Also hat die Projektion von P'' in die Ebene $x_1 = 0$ einen positiven ebenen Lebesgueschen Inhalt. Mithin ist $\mu_0(P'') > 0$. Wegen e) ist damit bewiesen: Für die Teilmenge P' von P existiert eine projektive Abbildung h , welche die Menge P' in eine beschränkte Menge P'' mit $\mu_K(P'') > 0$ transformiert. Indem wir nötigenfalls hinter h noch eine Ähnlichkeitstransformation schalten, können wir h sofort als dehnungslos annehmen. Dann ist aber wegen des Axioms c) von Kolmogoroff (S. 693) auch $\mu_K(P') > 0$ und wegen $P' \subset P$ auch $\mu_K(P) > 0$. Damit ist auch die letzte Ungleichung (2) bewiesen.

Aus (1) und (2) folgt jetzt mit Hilfe des Satzes 2 sofort der Satz 5.

Korollar. Die Maße μ_J , μ_0 und μ_F sind keine Kolmogoroff-Maße μ_K .

Satz 6. Es gibt im E_3 eine Fläche F , für welche $\mu_C(F) < \mu_H(F)$ ist.

Beweis. Nach A. S. Besicovitch¹⁸⁾ existiert in der x_1x_2 -Ebene des E_3 eine perfekte, nulldimensionale Menge Q , für welche $L_C(Q) < L_H(Q) = \frac{2}{\sqrt{3}}$ ist; dabei bedeuten L_C und L_H die linearen Maße von C. Carathéodory und F. Hausdorff; deren Definitionen lauten genau so wie die für μ_C und μ_H , nur treten an die Stelle der zweidimensionalen Durchmesser die gewöhnlichen Durchmesser.

Für eine Strecke S der x_3 -Achse mit der Länge s bezeichnen wir mit R^S die Menge aller Punkte (x_1, x_2, x_3) des E_3 , für welche (x_1, x_2) in Q und x_3 in S liegt. Wir behaupten die beiden Ungleichungen

$$(3) \quad \mu_C(R^S) \leq s \cdot L_C(Q), \quad \mu_H(R^S) \geq s \cdot L_H(Q).$$

Die erste Ungleichung ist richtig nach J. F. Randolph¹⁹⁾. Zum Beweis der zweiten Ungleichung wählen wir für eine beliebige Zahl $c < L_H(Q)$ ein $\delta > 0$ derart, daß für jede Überdeckung von Q durch endlich oder abzählbar unendlich viele Kreisscheiben U_i mit Durchmessern $\epsilon_i < \delta$ die Ungleichung $\sum \epsilon_i > c$ gilt. Nun seien V_1, V_2, \dots endlich oder abzählbar unendlich viele offene Kugeln mit Durchmessern $d_i < \delta$, welche R^S überdecken und zu R^S nicht fremd sind. Für ein beliebiges x_3 der Strecke S der x_3 -Achse bezeichnen wir mit $U_i(x_3)$ den Durchschnitt von V_i mit der zur x_1x_2 -Ebene parallelen Ebene E durch den Punkt x_3 . Ist $U_i(x_3)$ leer, so setzen wir $\epsilon_i(x_3) = 0$; andernfalls ist $U_i(x_3)$ eine offene Kreisscheibe mit einem Durchmesser $\epsilon_i(x_3) < \delta$. Da diese Kreisscheiben den zu Q kongruenten Durch-

¹⁸⁾ Math. Annalen 96 (1928), S. 458–464.

¹⁹⁾ J. F. Randolph, Bull. Amer. Math. Soc. XLII (1936), S. 269.

schnitt $E \cdot R^S$ überdecken, ist $\sum_i e_i(x_3) > c$ für jedes x_3 nach Wahl von δ .

Also ist, wenn a und b die Endpunkte der Strecke S sind,

$$\int_{a-\delta}^{b+\delta} \sum_i e_i(x_3) dx_3 \geq (b-a+2\delta) \cdot c = (s+2\delta) \cdot c \geq s \cdot c.$$

Nun ist aber

$$\int_{a-\delta}^{b+\delta} \sum_i e_i(x_3) dx_3 = \sum_i \int_{a-\delta}^{b+\delta} e_i(x_3) dx_3 = \frac{\pi}{4} \sum_i d_i^2.$$

Die Summe der zweidimensionalen Durchmesser $\frac{\pi}{4} d_i^2$ der Kugeln V_i ist also $\geq s \cdot c$. Nach Wahl der Zahl c und der Kugeln V_i folgt hieraus die zweite Ungleichung (3).

Nun sei L wieder ein Diskontinuum im Einheitsintervall E der x_3 -Achse mit positivem Lebesgueschem Inhalt $m^{(1)}L$. Dann kann man schreiben:

$L = \prod_{j=1}^{\infty} \sum_k S_{jk}$, wobei $\sum_k S_{jk}$ eine Summe von endlich vielen paarweise fremden Strecken S_{j1}, S_{j2}, \dots und jedes $S_{j+1,k}$ in einem S_{jk} enthalten ist. Dann ist $m^{(1)}L = \lim_{j \rightarrow \infty} \sum_k s_{jk}$, wobei s_{jk} die Länge von S_{jk} ist.

Es sei schließlich P die Menge aller Punkte (x_1, x_2, x_3) des E_3 mit $(x_1, x_2) \in Q$ und $x_3 \in L$. Dann ist $P = \prod_{j=1}^{\infty} \sum_k R^{S_{jk}}$ und

$$\mu_C(P) = \lim_j \sum_k \mu_C(R^{S_{jk}}), \quad \mu_H(P) = \lim_j \sum_k \mu_H(R^{S_{jk}}).$$

Wegen (3) ist also

$$\mu_C(P) \leq \lim_j \sum_k s_{jk} \cdot L_C(Q) = m^{(1)}L \cdot L_C(Q)$$

$$\mu_H(P) \geq \lim_j \sum_k s_{jk} \cdot L_H(Q) = m^{(1)}L \cdot L_H(Q).$$

Nun ist $m^{(1)}L > 0$ und $\mu_C(Q) < \mu_H(Q)$ nach Wahl von L und Q . Also ist $\mu_C(P) < \mu_H(P)$.

Hieraus folgt nach Satz 2 der Satz 6.

Zusatz. Für die Fläche F des Satzes 6 gilt $\mu_H(F) < \infty$.

Beweis. Nach Satz 2 genügt es zu zeigen, daß $\mu_H(P) < \infty$ ist. Wegen $P \subset R^E$, wobei E das Einheitsintervall der x_3 -Achse ist, genügt es weiter, $\mu_H(R^E) < \infty$ zu beweisen. Hierzu überdecken wir die Menge Q der $x_1 x_2$ -Ebene, deren lineares Hausdorff-Maß gleich $\frac{2}{\sqrt{3}}$ ist, durch endlich viele Kreise K_1, \dots, K_l , deren Durchmesser d_i kleiner als eine beliebig vorgegebene Zahl ϱ zwischen 0 und 1 sind und eine Summe $\sum_{i=1}^l d_i < \frac{4}{\sqrt{3}}$ haben. Betrachten

wir für jedes $\lambda = 1, \dots, l$ den geraden Vollzylinder Z_λ mit der Grundfläche K_λ und der Höhe 1. Er ist Summe von $\left[\frac{1}{d_\lambda}\right]$ Zylindern der Höhe d_λ und, falls $\frac{1}{d_\lambda}$ nicht ganz ist, einem weiteren Zylinder mit kleinerer Höhe. Jeder dieser höchstens $\frac{d_\lambda + 1}{d_\lambda}$ Zylinder mit der Summe Z_λ ist enthalten in einer Kugel vom Durchmesser $d_\lambda \sqrt{2} < \varrho \cdot \sqrt{2}$. Die Summe der zweidimensionalen Durchmesser dieser Kugeln ist $\leq \pi \frac{d_\lambda^2}{2} \cdot \frac{d_\lambda + 1}{d_\lambda} = \frac{\pi}{2} (d_\lambda + 1) d_\lambda \leq \frac{\pi}{2} (\varrho + 1) d_\lambda \leq \pi d_\lambda$. Summation über $\lambda = 1, \dots, l$ ergibt, daß wir R^E überdecken können durch endlich viele Kugeln mit Durchmessern $< \varrho \sqrt{2}$, deren zweidimensionale Durchmesser eine Summe $< \pi \frac{4}{\sqrt{3}}$ haben. Also ist auch $\mu_H(R^E) \leq \pi \frac{4}{\sqrt{3}}$, also endlich, wie zu zeigen war.

Satz 7. *Es gibt im E_3 eine Fläche F , für welche die Maße $\mu_K(F)$ nicht zusammenfallen.*

Beweis. An anderer Stelle²⁰⁾ haben wir gezeigt, daß für jedes dehnungslose Bild $A = f(A') \subset E_3$ einer ebenen analytischen Menge A' u. a. die Gleichung $\mu_C(A) = \mu_H(A)$ besteht. Also ist die Fläche F des Satzes 6 nicht darstellbar als dehnungsloses Bild einer ebenen analytischen Menge. Mithin ist das maximale Kolmogoroff-Maß²¹⁾ $\bar{\mu}_K(F) = \infty$. Andererseits ist $\mu_H(F) < \infty$ nach dem Zusatz zu Satz 6. Nun ist das Hausdorff-Maß μ_H ein Kolmogoroff-Maß μ_K ²²⁾. Also ist $\mu_K(F) < \bar{\mu}_K(F)$.

Eine Vermutung. A. S. Besicovitch²²⁾ hat eine ebene perfekte Menge Q angegeben, für welche das lineare Kolmogoroffsche Minimalmaß kleiner ist als das lineare Maß von Carathéodory. Es ist uns nicht gelungen, hieraus eine nulldimensionale kompakte Menge P des E_3 herzuleiten, für welche $\mu_K(P) < \mu_C(P)$ ist. Wir zweifeln jedoch nicht daran, daß es ein solches P gibt. Dann existiert nach Satz 2 auch eine Fläche F im E_3 mit der Eigenschaft $\mu_K(F) < \mu_C(F)$.

§ 4.

Schlußbemerkungen.

Wir wollen uns die Gründe für das Auseinanderfallen der von uns behandelten Maße klarzumachen versuchen. Betrachten wir zunächst das Maß μ_C von Carathéodory. Man kann die Definition von μ_C auch so fassen:

²⁰⁾ Erscheint demnächst in der Math. Zeitschr.

²¹⁾ $\bar{\mu}_K$ ist die untere Grenze der Lebesgueschen Inhalte $m^{(2)} A'$ der analytischen ebenen Mengen A' , die sich dehnungslos auf A abbilden lassen; existieren solche Mengen A' nicht, so wird $\bar{\mu}_K(A) = +\infty$ gesetzt.

²²⁾ Math. Annalen 113 (1937), S. 416–423.

Man überdecke die Menge M durch endlich oder abzählbar unendlich viele (nicht notwendig konvexe) Mengen U_i und bilde die Summe ihrer zweidimensionalen Durchmesser d_i , die folgendermaßen definiert sind: man projiziere U_i in eine Ebene E und bilde den Inhalt der konvexen Hülle dieser Projektion; die obere Grenze dieser Inhalte für alle Ebenen E ist der zweidimensionale Durchmesser von U_i ; im übrigen verfähre man wie in § 2, III. In dieser Fassung kann man das Maß μ_C ebenfalls als ein Projektionsmaß auffassen. Der wesentliche Unterschied dieses Projektionsmaßes vom Maß μ_G von Groß besteht darin, daß bei der Definition des zweidimensionalen Durchmessers nicht wie bei Groß das (äußere) Maß der Projektionen selbst, sondern ihrer konvexen Hüllen benutzt wird. Dies macht es verständlich, daß das Maß μ_G kleiner sein kann als das Maß μ_C und um so mehr kleiner als das Maß μ_H , das niemals kleiner sein kann als μ_C und für welches der analoge Unterschied gegenüber μ_G besteht.

Vergleichen wir nun die Maße μ_G und μ_F miteinander. Groß verwendet, wie gesagt, zwar die äußeren Maße der Projektionen selbst, aber er bestimmt die obere Grenze dieser Maße, benutzt also, kurz gesagt, nur die „Maximalprojektion“. Hingegen berücksichtigt Favard sämtliche Projektionen. Man kann nämlich das Favardsche Maß μ_F auch folgendermaßen definieren²⁰⁾: Für ein $\varepsilon > 0$ stelle man die Menge M dar als Summe von endlich oder abzählbar unendlich vielen paarweise fremden Teilmengen $M_1, M_2, \dots < \varepsilon$; es sei nun $\nu(M_i) = \frac{1}{\pi} \int m^{(2)} M_{iE} \dot{\varphi}$ gesetzt, wobei $m^{(2)} M_{iE}$ das (äußere) Lebesguesche Maß der Projektion M_{iE} von M_i in die Ebene E durch den Ursprung, $\dot{\varphi}$ das Flächenelement des sphärischen Normalenbildes der Ebenen E ist und integriert wird über alle Ebenen E durch den Ursprung; dann ist $\mu_F = \lim_{\varepsilon \rightarrow 0} \sum_i \nu(M_i)$. Es ist also nicht erstaunlich, daß μ_F von μ_G und um so mehr von μ_C und μ_H verschieden sein kann.

(Eingegangen am 19. 5. 1942.)

Über die Umordnungsstärke und eine Erweiterung des Steinitzschen Satzes.

Von

H. Hadwiger in Bern (Schweiz).

A. Einleitende Bemerkungen.

Es sei

$$(1) \quad \sum_{v=1}^{\infty} v_v = 0$$

eine *bedingt konvergente* Vektorreihe des s -dim. Vektorraumes. Die unwesentliche Voraussetzung, daß die Summe der Reihe Null ist, hat rein formale Bedeutung. Die Reihe kann bekanntlich so umgeordnet werden, daß die Konvergenz erhalten bleibt, jedoch andere Summen erzielt werden¹⁾. Die Gesamtheit aller als Summe konvergenter Umordnungen darstellbaren Vektoren nennen wir *Summenmenge* S der Reihe. Es gilt dann der bekannte Satz von E. Steinitz²⁾:

Die Summenmenge S einer bedingt konvergenten Vektorreihe ist eine lineare Vektormannigfaltigkeit.

In der vorliegenden Arbeit wird nun ein neuer Gesichtspunkt dadurch gewonnen, daß zunächst ein geeignetes Maß für die Stärke einer Umordnung eingeführt und dann festgestellt wird, daß zwischen der *Umordnungsstärke* und der entsprechenden Umordnungssumme eine allgemein faßbare Bindung besteht.

Es werden Umordnungen betrachtet, deren Stärke einen festen Wert α nicht erreichen; diese werden zu sogenannten *Stärkeklassen* U_α zusammengefaßt. Umordnungen, die zu U_1 gehören, nennen wir insbesondere *schwach*. Wir machen dann die Feststellung, daß die zu einer Stärkeklasse U_α ($\alpha \leq 1$) gehörende Summenmenge S_α für sich eine lineare Vektormannigfaltigkeit ist, wodurch der Satz von Steinitz in diesem Sinne eine Erweiterung erfährt (Satz II). Je nach der Natur der vorliegenden Reihe und der Größe von α wird S_α eine echte oder unechte Teilmannigfaltigkeit von S sein.

¹⁾ Über den Riemannschen Satz über Umordnung von Reihen gewöhnlicher Zahlen vgl. K. Knopp, *Theorie und Anwendung der unendlichen Reihen*, Berlin 1924, S. 319. Für Reihen komplexer Zahlen gab W. Threlfall, *Math. Zeitschr.* **24** (1926), S. 212–214, einen besonders kurzen Beweis dieser an sich fundamentalen Eigenschaft der bedingt konvergenten Reihen.

²⁾ E. Steinitz, *Journ. reine u. angew. Math.* **143** (1913), S. 128–175 (bes. S. 171, 172); W. Gross, *Monatshefte Math. u. Phys.* **28** (1917), S. 221–237; V. Bergström, *Abhandl. math. Seminar Hamburg* **8** (1931), S. 148–152; A. Wald, *Ergebnisse eines math. Kolloquiums Wien*, **59**, Koll. (22. II. 1933).

Besondere Bedeutung kommt den Fällen zu, wo S_α die triviale Mannigfaltigkeit Null (Nullvektor) ist. Die zu einer solchen Stärkeklasse U_α gehörenden Umordnungen sind wirkungslos, das heißt, sie ändern die Reihensumme nicht. Es zeigt sich, daß für die Charakterisierung derartiger Umordnungen eine durch die Beträge der Vektoren der vorgelegten Reihe bestimmte Zahl λ bedeutungsvoll wird. Der Konvergenzexponent λ der Reihe (1) ist wie folgt definiert:

Für $\beta > 0$ wird die Reihe

$$(2) \quad \sum_{v=1}^{\infty} |v_v|^\beta$$

entweder konvergieren oder divergieren. Wir setzen

$\lambda = \infty$, falls es keine Zahl β gibt, so daß die Reihe (2) konvergiert; andernfalls sei λ die untere Schranke aller β , für welche die Reihe (2) konvergent ausfällt. Falls λ endlich ist, gilt²⁾

$$(3) \quad \lambda = \limsup_{v \rightarrow \infty} - \frac{\log v}{\log |v_v|}.$$

Im Hinblick auf die Voraussetzung, wonach die Reihe (1) nur bedingt konvergiert, stellen wir fest, daß $\lambda \geq 1$ sein wird, da die Reihe (2) für $\beta = 1$ offenbar divergiert.

Wir beschränken uns in dieser Arbeit nur auf solche Fälle, die einen endlichen Konvergenzexponenten λ liefern.

In bezug auf die oben berührte Frage der wirkungslosen Umordnungen werden wir zeigen (Satz Ia), daß die Stärkeklasse U_α die triviale Vektormannigfaltigkeit Null ist, falls $\alpha \leq \frac{1}{\lambda}$ ist. Die angegebene Zahl $\frac{1}{\lambda}$ ist ferner die beste nur von λ abhängige Schranke für wirkungslose Umordnungen (Satz Ib). Das Beispiel am Schluß zeigt, daß durch Beschränkungen der Umordnungen auf Stärkeklassen tatsächlich echte Teilmannigfaltigkeiten aus der Steinitzischen Summenmenge S ausgeschieden werden können.

B. Definition der Umordnungsstärke.

Es bezeichne

$$(4) \quad \sum_{v=1}^{\infty} v_v'$$

eine Umordnung der Reihe (1). Zu jedem Index $n \geq 1$ lassen sich zwei Zahlen $N \geq 0$ und $M > 0$ so finden, daß

$$(5) \quad \sum_{v=1}^N v_v < \sum_{v=1}^n v_v' < \sum_{v=1}^M v_v$$

²⁾ Vgl. G. Pólya und G. Szegő, Aufgaben und Lehrsätze aus der Analysis. I. Berlin 1925; S. 19, 20, bes. Aufgabe Nr. 113.

gilt, wo das Zeichen \subset die Aussage darstellt, daß alle im Teilabschnitt links vorkommenden Vektoren auch rechts stehen. Der Teilabschnitt ganz links fällt weg, wenn $N = 0$ ist. Wir wählen nun von allen in Betracht fallenden Zahlen N die größte $N(n)$, und von allen M die kleinste $M(n)$ aus. Diese beiden eindeutig durch n bestimmten *Hüllzahlen* $N(n)$ und $M(n)$ bilden monoton nicht abnehmende Folgen. Offenbar gilt

$$(6) \quad 0 \leq N(n) \leq n \leq M(n).$$

Nun setzen wir

$$(7) \quad x_n = \frac{\log [1 + M(n) - N(n)]}{\log [2 + N(n)]}.$$

Die in Aussicht genommene Definition der Umordnungsstärke x der Umordnung (4) basiert auf der durch (7) festgelegten Folge. Wir definieren:

Es ist $x = \infty$, wenn die Folge (7) unbeschränkt ist; andernfalls ist

$$(8) \quad x = \limsup_{n \rightarrow \infty} x_n.$$

Unmittelbar aus der angegebenen Definition lassen sich einige einfache Folgerungen ziehen: Im Falle der identischen Umordnung ist $N(n) = M(n) = n$ für alle n , also $x_n = 0$; die Umordnungsstärke erhält den Wert 0. Das nämliche trifft bei allen Umordnungen zu, die nur in einem endlichen Abschnitt der Reihe wirksam sind. Es können ferner Umordnungen betrachtet werden bei welchen die Distanzen der Reihenglieder von ihren ursprünglichen Plätzen einen festen endlichen Wert nicht überschreiten. Die Differenzen $M(n) - N(n)$ bleiben dann beschränkt und die Umordnungsstärke erhält auch in diesen Fällen den Wert 0. Den Resultaten des folgenden Abschnitts kann entnommen werden, daß eine Umordnung der Stärke 0 die Summe der Reihe nicht verändert.

C. Wirkungslose Umordnungen.

Auf Grund der oben definierten Umordnungsstärke werden wir aus der Menge aller Umordnungen Teilmengen absondern, die wir *Stärkeklassen* nennen. Wir definieren: *Es sei $\alpha > 0$; die Stärkeklasse U_α bestehe aus der Gesamtheit aller Umordnungen, deren Umordnungsstärken $x < \alpha$ ausfallen.*

Besonders ausgezeichnet wird die Stärkeklasse U_1 ; sie umfaßt die *schwachen* Umordnungen.

U_0 bezeichne die Gesamtheit aller Umordnungen der Umordnungsstärke $x = 0$.

Eine Umordnung, die die Summe der Reihe nicht verändert, nennen wir *wirkungslos*.

Es ist naheliegend, die Frage aufzuwerfen, ob und wie sich wirkungslose Umordnungen durch die Umordnungsstärke charakterisieren lassen. Es ist

leicht einzusehen, daß es Umordnungen nicht verschwindender, aber beliebig kleiner Umordnungstärke gibt, die nicht wirkungslos sind; die erwähnte Charakterisierung kann also nicht unabhängig von der Reihe, die umgeordnet wird, vollzogen werden.

Es ist plausibel, daß dem asymptotischen Verhalten der Beträge der in der Reihe enthaltenen Vektoren wesentliche Bedeutung zukommt; wenn diese Beträge rasch gegen Null abnehmen, werden stärkere Umordnungen nötig sein, um einen bestimmten Umordnungseffekt zu erzielen, als wenn sie langsam gegen Null streben. Ein geeignetes Maß zur Kennzeichnung der Stärke der oben erwähnten Konvergenz der Vektorbeträge gegen Null ist der reziproke Wert des durch (3) definierten und nach Voraussetzung endlichen Konvergenz-exponenten λ .

Die beiden folgenden Sätze ergeben nun die exakte Klärung des oben geschilderten Sachverhaltes.

In den Formulierungen der Sätze verwenden wir den in A definierten Begriff der *Summenmenge*. Diese kann sinngemäß auch einer bestimmten Teilmenge von Umordnungen, z. B. einer Stärkeklasse zugewiesen werden; die Summenmenge einer Klasse von wirkungslosen Umordnungen ist Null, d. h. besteht nur aus dem Nullvektor. Die Summe der ursprünglichen Reihe (1), die umgeordnet wird, ist hier wie im folgenden stets als 0 vorausgesetzt.

Die beiden Sätze lauten:

Satz Ia. Die der Stärkeklasse U_α von Umordnungen einer Reihe vom Konvergenzexponent λ zugewiesene Summenmenge ist Null, falls $\alpha \leq \frac{1}{\lambda}$ ist.

Die Stärkegrenze $\frac{1}{\lambda}$ für wirkungslose Umordnungen ist in dem Sinne genau, als sie in Satz Ia nicht durch eine höhere ersetzt werden kann; d. h. es gilt folgende Umkehrung:

Satz Ib. Ist die Summenmenge der Umordnungen der Stärkeklasse U_α für jede Reihe vom Konvergenzexponenten λ Null, so ist $\alpha \leq \frac{1}{\lambda}$.

Im Hinblick auf $\lambda \geq 1$ ergibt sich, daß die durch Satz Ia gekennzeichneten wirkungslosen Umordnungen stets schwach sind. Wie bereits in B angekündigt, ergibt sich ferner, daß die der Stärkeklasse U_0 zugewiesene Summenmenge Null ist, daß also jede Umordnung verschwindender Stärke wirkungslos ist. Man beachte, daß die letzte Aussage unabhängig von der Reihe gemacht wird.

Beweis von Satz Ia:

Es bezeichne

$$(9) \quad \sum_{r=1}^{\infty} v_r' = w$$

eine konvergente Umordnung der Reihe (1), der auf Grund von (3) der endliche Konvergenzexponent λ zukommt. Die durch (8) definierte Umordnungsstärke von (9) sei x , und es bestehe die Beziehung $x < \frac{1}{\lambda}$. Es ist zu zeigen, daß die Summe w der Reihe (9) verschwinden muß. Nach den in B gegebenen Definitionen kann die n -te Teilsumme von (9) geschrieben werden:

$$(10) \quad \sum_{v=1}^n v_r' = \sum_{v=1}^{N(n)} v_r + \sum_{v=1}^{\omega_n} v_{\theta_v},$$

wo $N(n) < \varrho_v \leq M(n)$, also

$$(11) \quad \omega_n \leq M(n) - N(n)$$

ist. Mit Rücksicht auf Definition (8) für x und auf die Voraussetzung muß

$$\limsup_{n \rightarrow \infty} \frac{\log [1 + \omega_n]}{\log [2 + N(n)]} < \frac{1}{\lambda}$$

ausfallen, und wegen $N(n) \rightarrow \infty$ kann ein θ , $0 < \theta < 1$, und ein n_0 so angegeben werden, daß für alle $n > n_0$

$$(12) \quad \frac{\log \omega_n}{\log N(n)} < \frac{1-\theta}{\lambda}, \quad \text{also} \quad \omega_n < N(n)^{\left(\frac{1-\theta}{\lambda}\right)}$$

ist. Setzen wir weiter

$$(13) \quad \lambda_v = - \frac{\log v}{\log |v_r|},$$

so kann wegen $\lambda = \limsup_{v \rightarrow \infty} \lambda_v$ [Definition (3)] ein v_0 so bestimmt werden, daß für alle $v > v_0$

$$(14) \quad \frac{1}{\lambda_v} - \frac{1}{\lambda} > - \frac{\theta}{2\lambda}$$

ausfällt; damit wird für $v > v_0$

$$\log |v_r| < \left(\frac{\theta^*}{2\lambda} - \frac{1}{\lambda} \right) \log v$$

und

$$(15) \quad |v_r| < v^{-\left(\frac{1}{\lambda} - \frac{\theta}{2\lambda}\right)} = v^{-\frac{\theta}{2\lambda}} v^{-\left(\frac{1-\theta}{\lambda}\right)}.$$

Wir wählen nun ein n_1 so, daß für alle $v > N(n)$, $n > n_1$,

$$(16) \quad |v_r| < \frac{\varepsilon}{2} v^{-\left(\frac{1-\theta}{\lambda}\right)}$$

ist. Endlich kann wegen der Konvergenz der Reihe (1) ein n_2 so festgelegt werden, daß für alle $n > n_2$

$$(17) \quad \left| \sum_{v=1}^{N(n)} v_r \right| < \frac{\varepsilon}{2}$$

bleibt. Nach diesen Vorbereitungen kehren wir zu der Zerlegung (10) zurück. Es bezeichne N die größte der drei Zahlen n_0, n_1, n_2 . Für jedes $n > N$ gelten dann gleichzeitig die Relationen (12), (16) und (17). Man erzielt zunächst

$$\left| \sum_{r=1}^n v_r' \right| \leq \left| \sum_{r=1}^{N(n)} v_r \right| + \sum_{r=1}^{n_n} |v_{q_r}|.$$

Da $q_r > N(n)$ ist, kommt (16) zur Anwendung, und es folgt wegen

$$|v_{q_r}| < \frac{\varepsilon}{2} q_r^{-\left(\frac{1-\sigma}{1}\right)} < \frac{\varepsilon}{2} N(n)^{-\left(\frac{1-\sigma}{1}\right)}$$

vorherst

$$\left| \sum_{r=1}^n v_r' \right| < \frac{\varepsilon}{2} \left[1 + \omega_n N(n)^{-\left(\frac{1-\sigma}{1}\right)} \right],$$

und mit Rücksicht auf (12) endlich

$$\left| \sum_{r=1}^n v_r' \right| < \varepsilon.$$

Hieraus folgt $|w| \leq \varepsilon$, und also die Behauptung $w = 0$.

Beweis von Satz Ib: Es genügt offenbar zu zeigen, daß zu jedem $\lambda \geq 1$ eine nicht wirkungslose Umordnung einer Reihe vom Konvergenzexponenten λ der Umordnungsstärke $\frac{1}{\lambda}$ existiert.

Es sei e ein Einheitsvektor und $\sigma = \frac{1}{\lambda}$. Wir betrachten die Reihe

$$(18) \quad \sum_{r=1}^{\infty} v_r = e - e + \left(\frac{1}{2}\right)^{\sigma} e - \left(\frac{1}{2}\right)^{\sigma} e + \left(\frac{1}{3}\right)^{\sigma} e - \left(\frac{1}{3}\right)^{\sigma} e + \dots = 0.$$

Der ihr zukommende Konvergenzexponent ist λ .

1. Fall: $\lambda > 1$. Durch

$$(19) \quad \begin{aligned} A(n) &= n + [n^{\sigma}] \\ B(n) &= n - [n^{\sigma}] \end{aligned} \quad (n = 1, 2, 3, \dots)$$

sind zwei monoton nicht abnehmende Zahlenfolgen definiert. Wegen

$$A(n) + B(n) = 2n$$

kann

$$(20) \quad \sum_{r=1}^{A(n)} \left(\frac{1}{r}\right)^{\sigma} e - \sum_{r=1}^{B(n)} \left(\frac{1}{r}\right)^{\sigma} e$$

als $2n$ -te Teilsumme einer Umordnung $\sum_{r=1}^{\infty} v_r'$ der Reihe (18) aufgefaßt werden.

Die dieser Umordnung zugeordneten Hüllzahlen können für geradzahliges Argument auf Grund von (20) sofort angeschrieben werden; es gilt

$$(21) \quad \begin{aligned} N(2n) &= 2B(n) + 1 = 2n + 1 - 2[n^{\sigma}], \\ M(2n) &= 2A(n) - 1 = 2n - 1 + 2[n^{\sigma}]. \end{aligned}$$

Im Hinblick auf das monotone Verhalten der Hüllzahlen genügt (21), um die Umordnungsstärke nach (7) und (8) zu ermitteln.

Es ergibt sich

$$(22) \quad x = \lim_{n \rightarrow \infty} \frac{\log [1 + M(n) - N(n)]}{\log [2 + N(n)]} = \sigma.$$

Für die 2-n-te Teilsumme der umgeordneten Reihe erhält man nach (20)

$$\sum_{v=1}^{2n} v'_v = \varrho_n e,$$

wo

$$\varrho_n = \frac{1}{(1+n-[n^a])^a} + \dots + \frac{1}{(n+[n^a])^a}$$

ist. Die einfache Abschätzung

$$\frac{2[n^a]}{(1+n+[n^a])^a} < \varrho_n \leq \frac{2[n^a]}{(1+n-[n^a])^a}$$

gestattet, auf

$$\lim_{n \rightarrow \infty} \varrho_n = 2$$

zu schließen. Die Umordnung konvergiert und es gilt

$$(23) \quad \sum_{v=1}^{\infty} v'_v = 2e.$$

2. Fall: $\lambda = 1$. Wir setzen

$$(24) \quad \begin{aligned} A(n) &= 2n \\ B(n) &= n \end{aligned} \quad (n = 1, 2, 3, \dots)$$

und es läßt sich wegen $A(n) + B(n) = 3n$

$$(25) \quad \sum_{v=1}^{A(n)} \frac{1}{v} e - \sum_{v=1}^{B(n)} \frac{1}{v} e$$

als 3-n-te Teilsumme einer Umordnung $\sum_{v=1}^{\infty} v'_v$ der Reihe (18) auffassen. Für die zugeordneten Hüllzahlen findet sich

$$(26) \quad \begin{aligned} N(3n) &= 2B(n) + 1 = 2n + 1, \\ M(3n) &= 2A(n) - 1 = 4n - 1. \end{aligned}$$

Wieder genügt (26) zur Berechnung der Umordnungsstärke x , und man erzielt

$$(27) \quad x = \lim_{n \rightarrow \infty} \frac{\log [1 + M(n) - N(n)]}{\log [2 + N(n)]} = 1.$$

Für die 3-n-te Teilsumme der betrachteten Umordnung hat man

$$\sum_{v=1}^{3n} v'_v = \varrho_n e, \quad \varrho_n = \frac{1}{n+1} + \dots + \frac{1}{2n},$$

und da

$$\lim_{n \rightarrow \infty} \varrho_n = \log 2$$

ist, konvergiert die Umordnung, und zwar ist

$$(28) \quad \sum_{v=1}^{\infty} v'_v = \log 2 \cdot e.$$

Mit den in den beiden Fällen nachgewiesenen nicht wirkungslosen Umordnungen ist der gewünschte Nachweis erbracht.

D. Eine Erweiterung des Steinitzischen Satzes.

Unterwerfen wir die bedingt konvergente Vektorreihe (1) der Gesamtheit aller Umordnungen, die noch eine konvergente Reihe liefern, so ist die erzeugte Summenmenge S nach dem Satz von Steinitz eine lineare Mannigfaltigkeit der Dimension k ($1 \leq k \leq s$).

Wir denken uns nun nur die Umordnungen einer gewissen Stärkeklasse U_α auf die Reihe ausgeübt. Die so gewonnene Summenmenge S_α wird offenbar eine echte oder unechte Teilmenge von S sein. Wir werden nun zeigen, daß für $\alpha \leq 1$, das heißt für eine schwache Klasse, S_α ebenfalls eine lineare Mannigfaltigkeit ist. Die Erweiterung des Steinitzischen Satzes besteht also in folgendem:

Satz II. *Die Summenmenge der Umordnungen der Stärkeklasse U_α ($\alpha \leq 1$) ist eine lineare Mannigfaltigkeit.*

Durch Satz Ia erfährt dieser Satz insofern eine Bestätigung, als dort für $\alpha \leq \frac{1}{\lambda}$ die Summenmenge die triviale lineare Mannigfaltigkeit Null darstellt.

Beweis von Satz II: Wir betrachten eine Vektormenge T_α ($\alpha \leq 1$), von der wir zeigen, daß sie

1. eine lineare Mannigfaltigkeit ist (1. Hilfsatz), und
2. mit der Summenmenge S_α zusammenfällt (2. Hilfsatz).

Wir charakterisieren T_α wie folgt:

Ein Vektor a gehört zu T_α , wenn sich eine Folge natürlicher Zahlen

$$(29) \quad \varrho(n) \geq 1 \quad (n = 0, 1, 2, \dots)$$

so angeben läßt, daß aus dem Teilstück

$$(30) \quad v_{n+1} + v_{n+2} + \dots + v_{n+\varrho(n)}$$

der Reihe (1) eine Auswahlsumme

$$(31) \quad a_n = \sum_{v=1}^{\varrho(n)} v_v,$$

gebildet werden kann, so daß

$$(32) \quad a = \lim_{n \rightarrow \infty} a_n$$

gilt, und außerdem

$$(33) \quad \limsup_{n \rightarrow \infty} \frac{\log \varrho(n)}{\log n} < \alpha$$

ausfällt.

1. Hilfssatz. Die Vektormenge T_a ist eine lineare Mannigfaltigkeit.

Wählt man aus dem Teilstück (30) die zu (31) komplementäre Auswahl

$$a'_n = \sum_{r=1}^{\varrho'_n} v_{n+r},$$

so daß

$$a_n + a'_n = v_{n+1} + v_{n+2} + \dots + v_{n+\varrho(n)},$$

so folgt im Hinblick auf die Konvergenz der ursprünglichen Reihe (1)

$$(34) \quad \lim_{n \rightarrow \infty} a'_n = - \lim_{n \rightarrow \infty} a_n = -a,$$

so daß erkannt wird, daß mit a auch $a' = -a$ zu T_a gehören muß. Es ist noch zu zeigen:

(a) Gehören a' und a'' zu T_a , so gehört auch $a' + a''$ zu T_a .

(b) Gehört a zu T_a , so gehört auch λa ($0 < \lambda < 1$) zu T_a .

Beweis für (a): Es gibt zwei Folgen natürlicher Zahlen $\varrho'(n)$ und $\varrho''(n)$, so daß

$$(35) \quad a'_n = \sum_{r=1}^{\varrho'_n} v_{n+r} \subset \sum_{r=1}^{\varrho'(n)} v_{n+r},$$

$$(36) \quad a''_m = \sum_{r=1}^{\varrho''_m} v_{m+r} \subset \sum_{r=1}^{\varrho''(m)} v_{m+r},$$

$$(37) \quad \lim_{n \rightarrow \infty} a'_n = a', \quad \lim_{m \rightarrow \infty} a''_m = a'',$$

$$(38) \quad \limsup_{n \rightarrow \infty} \frac{\log \varrho'(n)}{\log n} < \alpha, \quad \limsup_{m \rightarrow \infty} \frac{\log \varrho''(m)}{\log m} < \alpha$$

gilt. Setzen wir in (35)

$$m = n + \varrho'(n),$$

so sind die Auswahlsummen a'_n und a''_m vektorfremd. Die Vereinigungssumme

$$a''' = a'_n + a''_{n+\varrho'(n)}$$

läßt sich als Auswahlsumme

$$(39) \quad a''' \subset \sum_{r=1}^{\varrho'''(n)} v_{n+r}$$

auffassen, wo

$$\varrho'''(n) = \varrho'(n) + \varrho''[n + \varrho'(n)]$$

gesetzt werden muß. Offenbar gilt

$$(39) \quad \lim_{n \rightarrow \infty} a_n''' = a' + a''.$$

Wir haben noch nachzuweisen, daß die Bedingung (33) erfüllt ist. Es kann ein $\theta > 0$ so angegeben werden, daß für alle $n, m > N$

$$\frac{\log \varrho'(n)}{\log n} < \alpha - \theta \quad \text{und} \quad \frac{\log \varrho''(m)}{\log m} < \alpha - \theta \quad \text{und} \quad 0 < \alpha - \theta$$

ausfällt, so daß unter Beachtung von $0 < \alpha - \theta < 1$

$$\varrho'''(n) < n^{\alpha-\theta} + (n + n^{\alpha-\theta})^{\alpha-\theta} < 3n^{\alpha-\theta}$$

folgt. Daher gilt

$$(40) \quad \limsup_{n \rightarrow \infty} \frac{\log \varrho'''(n)}{\log n} \leq \alpha - \theta < \alpha.$$

Mit (38), (39) und (40) ist die Aussage (a) bewiesen.

Beweis für (b): Es sei

$$(41) \quad a_n = \sum_{\nu=1}^{\omega_n} v_{n\nu} \subset \sum_{\nu=1}^{\varrho(n)} v_{n\nu},$$

und

$$(42) \quad \lim_{n \rightarrow \infty} a_n = a.$$

Wir zerlegen $v_{n\nu}$ in zwei orthogonale Komponenten

$$(43) \quad w_{n\nu} = a_{n\nu} + w_{n\nu}, \quad a_{n\nu} \perp w_{n\nu},$$

wobei

$$(44) \quad a_{n\nu} = \gamma_{n\nu} a_n$$

sein soll. Es gilt dann

$$(45) \quad \sum_{\nu=1}^{\omega_n} w_{n\nu} = 0,$$

$$(46) \quad \sum_{\nu=1}^{\omega_n} \gamma_{n\nu} = 1.$$

Wir denken uns nun die Anordnung der Vektoren in der Auswahlsumme (41) so gewählt, daß das Maximum der Beträge der Teilsummen von (45) einen möglichst kleinen Wert annimmt. Nach einem Lemma von Gross und Bergström⁴⁾ ist dann

$$(47) \quad \left| \sum_{\nu=1}^n w_{n\nu} \right| \leq A q_n \quad (\mu = 1, 2, \dots, \omega_n),$$

⁴⁾ Vgl. die in Fußnote ²⁾ zitierten Arbeiten.

wo A eine nur von der Dimension s des Vektorraumes abhängige Konstante ist und q_n die größte der Zahlen $|w_{n1}|, |w_{n2}|, \dots, |w_{nc_n}|$ bedeutet. Wegen $|w_{nv}| \leq |v_{nv}|$ folgt mit Rücksicht auf die Konvergenz der Reihe (1)

$$(48) \quad \lim_{n \rightarrow \infty} q_n = 0.$$

Wenn ferner p_n die größte der Zahlen

$$|a_{n1}|, |a_{n2}|, \dots, |a_{nc_n}|$$

bedeutet, so folgt analog

$$(49) \quad \lim_{n \rightarrow \infty} p_n = 0.$$

Nun ist weiter

$$(50) \quad |\gamma_{nv}| = \frac{|a_{nv}|}{|a_n|} \leq \frac{p_n}{|a_n|}.$$

Zu einem gegebenen λ , $0 < \lambda < 1$, kann im Hinblick auf (46) und (50) ein σ_n so gewählt werden, daß

$$(51) \quad \lambda \leq \sum_{v=1}^{c_n} \gamma_{nv} \leq \lambda + \frac{p_n}{|a_n|}$$

gilt. Entsprechend bilden wir die Auswahlsumme

$$(52) \quad a'_n = \sum_{v=1}^{c_n} v_{nv} \subset \sum_{v=1}^{c(n)} v_{nv}.$$

Dann ist

$$a'_n - \lambda a_n = \left(\sum_{v=1}^{c_n} \gamma_{nv} - \lambda \right) a_n + \sum_{v=1}^{c_n} w_{nv},$$

so daß mit Rücksicht auf (47) und (51)

$$(53) \quad |a'_n - \lambda a_n| \leq p_n + A q_n,$$

und also nach (48), (49) und (42)

$$(54) \quad \lim_{n \rightarrow \infty} a'_n = \lambda a$$

folgt. Da die Auswahlsummen a'_n den gleichen Teilstücken zugeordnet sind wie die Auswahlsummen a_n , so ist die Bedingung (33) nach Voraussetzung erfüllt.

Damit ist der 1. Hilfssatz bewiesen.

2. Hilfssatz. Die Vektormenge T_a ist identisch mit der Summenmenge S_a .

Wir haben folgendes zu zeigen:

- (a) Gehört a zu S_a , so gehört a auch zu T_a .
- (b) Gehört a zu T_a , so gehört a auch zu S_a .

Beweis von (a): Es gibt eine Umordnung

$$(55) \quad a = \sum_{r=1}^{\infty} v_r$$

der Reihe (1) der Stärke $x < \alpha$. Zu jedem n kann ein l so angegeben werden, daß

$$(56) \quad \sum_{r=1}^n v_r \subset \sum_{r=1}^l v_r$$

gilt. Wählen wir bei festem n unter allen l die kleinste in Betracht fallende Zahl und bezeichnen diese mit $L(n)$, so ist

$$(57) \quad \sum_{r=1}^{L(n)} v_r = \sum_{r=1}^n v_r + \sum_{r=1}^{\omega n} v_{n+r},$$

wobei

$$(58) \quad a_n' = \sum_{r=1}^{\omega n} v_{n+r} \subset \sum_{r=n+1}^{M[L(n)]} v_r = \sum_{r=1}^{\varrho(n)} v_{n+r}$$

mit

$$(59) \quad \varrho(n) = M[L(n)] - n$$

gesetzt werden kann. Der Vektor a_n' ist damit als Auswahlsumme dargestellt. Aus der Konvergenz der Reihen (1) und (55) folgt in Verbindung mit der Zerlegung (57) bzw. (58) also

$$(60) \quad \lim_{n \rightarrow \infty} a_n' = a,$$

so daß die Bedingungen (31) und (32) für die Zugehörigkeit von a zu T_α erfüllt sind. Zeigen wir noch, daß auch die Bedingung (33) erfüllt ist:

Aus der Definition der Hüllzahlen $N(k)$ und $M(k)$ kann auf die Relation

$$(61) \quad N(k) \leq M(k-1) + 1$$

und aus der Konstruktion der Zahlen $L(n)$ auf

$$(62) \quad N[L(n)-1] < n$$

geschlossen werden. Mit Hilfe von (61) und (62) verifiziert man leicht

$$(63) \quad \begin{aligned} \varrho(n) &< M[L(n)] - N[L(n)] + M[L(n)-1] - N[L(n)-1] + 1, \\ n &> N[L(n)] - M[L(n)-1] + N[L(n)-1] - 1. \end{aligned}$$

Wegen der Voraussetzung über die Umordnung (55) gilt

$$(64) \quad \limsup_{k \rightarrow \infty} \frac{\log[1 + M(k) - N(k)]}{\log[2 + N(k)]} = x < \alpha,$$

so daß ein $\theta > 0$ so angegeben werden kann, daß für alle $k \geq L(n)-1$, wobei $n > N_0$ gewählt ist,

$$(65) \quad M(k) - N(k) < \{N(k)\}^{\alpha-\theta}$$

ausfällt. Verwenden wir diese Abschätzung für (63), so ergibt sich

$$(66) \quad \begin{aligned} \varrho(n) &< 2 \{N[L(n)]\}^{\alpha-\theta} + 1, \\ n &> N[L(n)] - \{N[L(n)]\}^{\alpha-\theta} - 1, \end{aligned}$$

so daß man mit Rücksicht auf $L(n) \rightarrow \infty$ und $0 < \alpha - \theta < 1$

$$(67) \quad \limsup_{n \rightarrow \infty} \frac{\log \varrho(n)}{\log n} \leq \alpha - \theta < \alpha$$

schließen kann, womit dargetan ist, daß auch die Bedingung (33) erfüllt ist.

Beweis von (b): Es sei

$$(68) \quad a = \lim_{n \rightarrow \infty} a_n,$$

$$(69) \quad a_n = \sum_{v=1}^{p_n} v_n, \subset \sum_{v=1}^{\varrho(n)} v_{n+v},$$

wobei

$$(70) \quad \limsup_{n \rightarrow \infty} \frac{\log \varrho(n)}{\log n} < \alpha$$

ist. Wir bilden nun die folgenden Summen

$$(71) \quad \mathfrak{S}_k = \sum_{v=1}^{p_k} v_v + a_{p_k},$$

wo

$$(72) \quad \begin{aligned} p_0 &= \varrho(0), \\ p_k &= p_{k-1} + \varrho(p_{k-1}) \end{aligned}$$

gesetzt wird. Mit Rücksicht auf (69) gilt offenbar

$$(73) \quad \mathfrak{S}_k \subset \mathfrak{S}_{k+1},$$

so daß \mathfrak{S}_{k+1} aus \mathfrak{S}_k dadurch erhalten werden kann, daß eine gewisse Anzahl von Vektoren addiert werden. Dies läßt sich auch so ausdrücken:

$$(74) \quad \mathfrak{S}_{k+1} - \mathfrak{S}_k = \sum_{v=1}^{q_k} v_{k+v}.$$

Wir denken uns nun die Reihenfolge der Vektoren in der Summe (74) so gewählt, daß das Maximum der Beträge der Teilsummen möglichst klein wird. Nach dem schon früher angewendeten Theorem von Gross und Bergström gilt dann

$$(75) \quad \left| \mathfrak{S}_{k+1} - \mathfrak{S}_k - \sum_{v=1}^s v'_v \right| \leq A a_k,$$

wo a_k die größte der Zahlen

$$|v'_{k1}|, |v'_{k2}|, \dots, |v'_{ks}|; |\mathfrak{S}_{k+1} - \mathfrak{S}_k|$$

und A eine nur von der Dimension s des Vektorraumes abhängige Zahl bezeichnet.

Aus (71) folgt man im Hinblick auf die Konvergenz der Reihen (1) und (68) mit $k \rightarrow \infty$, also auch $p_k \rightarrow \infty$,

$$(76) \quad \lim_{k \rightarrow \infty} \mathfrak{S}_k = a,$$

so daß

$$(77) \quad \lim_{k \rightarrow \infty} a_k = 0$$

gilt. Wir bilden nun die Reihe

$$(78) \quad \mathfrak{S}_0 + \sum_{r=1}^{a_0} v'_0 + \sum_{r=1}^{a_1} v'_1 + \dots = \sum_{r=1}^{\infty} v'_r,$$

die offenbar eine Umordnung der Reihe (1) darstellt. Zu einem natürlichen

$$l > \varrho(0) + \varrho[\varrho(0)]$$

kann ein k und ein τ so gegeben werden, daß

$$\sum_{r=1}^l v'_r = \mathfrak{S}_k + \sum_{r=1}^{\tau} v'_r,$$

und somit

$$\mathfrak{S}_{k+1} - \sum_{r=1}^l v'_r = \mathfrak{S}_{k+1} - \mathfrak{S}_k - \sum_{r=1}^{\tau} v'_r,$$

oder wegen (75)

$$\left| \mathfrak{S}_{k+1} - \sum_{r=1}^l v'_r \right| \leq A a_k$$

gilt. Mit Rücksicht auf (76) und (77) folgt

$$(79) \quad \lim_{l \rightarrow \infty} \sum_{r=1}^l v'_r = a,$$

da aus $l \rightarrow \infty$ notwendig auch $k \rightarrow \infty$ folgt. Wir haben somit eine konvergente Umordnung

$$(80) \quad a = \sum_{r=1}^{\infty} v'_r$$

der Reihe (1) mit der Summe a gefunden. Es ist noch zu zeigen, daß ihre Umordnungstärke $x < \alpha$ ausfällt.

Aus (79) folgt mit Rücksicht auf die Konstruktion (71)

$$\mathfrak{S}_k \subset \sum_{r=1}^l v'_r \subset \mathfrak{S}_{k+1}.$$

Für die der Umordnung (80) zugewiesenen Hüllzahlen ergeben sich hieraus die Ungleichungen

$$(81) \quad N(l) \geq p_k, \quad M(l) \leq p_{k+2}.$$

Der Rekursion (72) entnehmen wir

$$(82) \quad p_{k+2} = p_{k+1} + e(p_{k+1}) = p_k + e(p_k) + e[p_k + e(p_k)].$$

Nach Voraussetzung (70) läßt sich ein $\theta > 0$ so angeben, daß für alle $l > L$

$$e(p_k) < \{p_k\}^{\alpha-\theta},$$

also nach (82)

$$p_{k+2} < p_k + \{p_k\}^{\alpha-\theta} + [p_k + \{p_k\}^{\alpha-\theta}]^{\alpha-\theta},$$

und wegen $0 < \alpha - \theta < 1$

$$(83) \quad p_{k+2} < p_k + 3 \{p_k\}^{\alpha-\theta}$$

ausfällt. Aus (81) folgert man nun in Verbindung mit (83)

$$\frac{\log [1 + M(l) - N(l)]}{\log [2 + N(l)]} < \frac{\log (1 + 3 \{p_k\}^{\alpha-\theta})}{\log (2 + p_k)} \quad \text{für } l > L,$$

und hieraus

$$(84) \quad \limsup_{l \rightarrow \infty} \frac{\log [1 + M(l) - N(l)]}{\log [2 + N(l)]} \leq \alpha - \theta,$$

oder also $x < \alpha$, was noch zu zeigen war. Damit ist auch der 2. Hilfssatz bewiesen.

E. Ein Beispiel.

Durch ein einfaches Beispiel wollen wir noch belegen, daß durch Beschränkungen der Umordnungstärke im Sinne von Satz II tatsächlich nicht triviale Teilmannigfaltigkeiten der Steinitzschen, Summenmannigfaltigkeit ausgeschieden werden können.

Wir betrachten eine Reihe komplexer Zahlen (Vektorreihe des R_2)

$$(85) \quad \sum_{v=1}^{\infty} v_v = 0,$$

wo

$$(86) \quad \begin{cases} v_{4\mu-3} = \frac{1}{\sqrt{\mu}} + \frac{i}{\sqrt{\mu}} \\ v_{4\mu-2} = -\frac{1}{\sqrt{\mu}} - \frac{i}{\sqrt{\mu}} \\ v_{4\mu-1} = \frac{1}{\sqrt{\mu}} - \frac{i}{\sqrt{\mu}} \\ v_{4\mu} = -\frac{1}{\sqrt{\mu}} + \frac{i}{\sqrt{\mu}} \end{cases} \quad (\mu = 1, 2, 3, \dots).$$

Bezeichnet w , den Imaginärteil von v , so gilt offenbar auch

$$(87) \quad \sum_{r=1}^{\infty} w_r = 0.$$

Für den der Reihe (85) bzw. (87) auf Grund von (3) zugeordneten Konvergenz-exponenten λ bzw. λ' erhält man den Wert

$$(88) \quad \lambda = 3 \text{ bzw. } \lambda' = 2.$$

In Verbindung mit der Feststellung betreffend die beim Beweis von Satz Ib behandelten speziellen Reihen folgt aus den Sätzen Ia und II der folgende Sachverhalt:

Ist S_a die Summenmenge der Umordnungen der Stärkeklasse der Reihe (85), so gilt:

$$\begin{array}{lll} 0 \leq \alpha \leq \frac{1}{2}: & S_a \equiv \{0\} & (\text{Punkt}) \\ \frac{1}{2} < \alpha \leq \frac{1}{2}: & S_a \equiv \{a\} & (a \text{ beliebig reell}) \quad (\text{Gerade}) \\ \frac{1}{2} < \alpha \leq 1: & S_a \equiv \{a + ib\} & (a, b \text{ beliebig reell}) \quad (\text{Ebene}). \end{array}$$

(Eingegangen am 6. 2. 1942.)

Modell einer Leibnizischen Differentialrechnung mit aktuell unendlich kleinen Größen sämtlicher Ordnungen.

Von

Ludwig Neder in Münster i. W.

Das Ziel der vorliegenden Arbeit ist der Nachweis der Widerspruchsfreiheit der Leibnizischen Differentialrechnung¹⁾ mit aktuell unendlich kleinen Größen (Differentialen) aller Ordnungen der Folge 1, 2, 3, ...

Die Grundlage unserer Betrachtungen ist ein komplexes Zahlssystem mit unendlich vielen Einheiten, das gedeutet werden kann als (eindimensionales) reelles System von gemischten Größen, die sich aus endlichen Größen und unendlich kleinen Größen 1., 2., 3., ... Ordnung additiv zusammensetzen (vgl. § 1, Nr. 3).

Von Interesse und Wichtigkeit ist ferner der hier benutzte (relativ komplizierte) Begriff der gemischten Funktion einer gemischten Veränderlichen. Derselbe stützt sich auf den folgenden Ausdruck des ersten Differentials $dF(x)$ der Funktion durch eine „Differentialpotenzreihe“

$$\sum_{r=1}^{\infty} r!^{-1} f^{(r)}(x) dx^r,$$

was als der wesentliche Gedanke der vorliegenden Arbeit zu betrachten ist.

Das vorstehend angeführte Zahlssystem ist nicht neu. Systeme, die auf es hinauslaufen oder noch allgemeiner sind, haben zu geometrischen Zwecken bereits Verwendung gefunden durch Veronese und andere Autoren²⁾. Dagegen scheint sich an den zugehörigen Funktionsbegriff noch niemand herangewagt zu haben.

§ 1.

Ein komplexes Zahlssystem mit unendlich vielen Einheiten.

1. *Definitionen.* Wir betrachten ein System von komplexen Zahlen der Gestalt

$$\left. \begin{aligned} A &= [a_0, a_1, a_2, \dots] = [a_v], a_v \text{ reell} \\ B &= [b_0, b_1, b_2, \dots] = [b_v], b_v \text{ reell} \\ X &= [x_0, x_1, x_2, \dots] = [x_v], x_v \text{ reell} \end{aligned} \right\} v = 0, 1, 2, \dots,$$

¹⁾ D. Mahnke, Neue Einblicke in die Entdeckungsgeschichte der höheren Analysis. Abh. Preuß. Akad. d. Wiss., Jahrg. 1925; Phys.-Math. Kl. Nr. 1.

²⁾ Vgl. die deutsche Enzyklopädie, Teil I, Art. A 5, Nr. 17 u. 18, und Teil III, Band 1, S. 119ff.

die wir als *gemischte* Größen bezeichnen. Eine gemischte Größe X heißt

endlich, wenn $x_0 \neq 0$;

unendlich klein, wenn $x_0 = 0$ aber nicht alle übrigen $x_\varrho = 0$;

Null, wenn $0 = x_0 = x_1 = x_2 = \dots$.

Eine gemischte Größe ist daher entweder endlich oder unendlich klein oder Null, und diese drei Fälle schließen sich gegenseitig aus. Ferner bezeichnen wir X als

unendlich klein der Ordnung ξ , wenn

$$x_0 = x_1 = \dots = x_{\xi-1} = 0 \quad \text{und} \quad x_\xi \neq 0.$$

Dabei zählen die endlichen Größen als unendlich klein der Ordnung 0.

2. Die *Gleichheit* ist erklärt durch die Formel

$$A = B \Leftrightarrow a_\varrho = b_\varrho \quad (\varrho = 0, 1, 2, \dots)$$

als reflexive, symmetrische und transitive Beziehung; und es ist die Null wohlverschieden von den unendlich kleinen Größen.

3. Die *Addition* ist erklärt durch die Formel

$$A + B = [a_\varrho + b_\varrho], \quad (\varrho = 0, 1, 2, \dots),$$

als eine assoziative und kommutative Verknüpfung mit dem Modul (Null-element)

$$O = [0, 0, 0, \dots].$$

Hiernach ist jede gemischte Größe X darstellbar in der Gestalt

$$X = [x_0, 0, 0, 0, \dots] + [0, x_1, x_2, x_3, \dots]$$

als Summe einer endlichen (oder verschwindenden) und einer unendlich kleinen (oder verschwindenden) Größe.

Ferner gilt: Die Summe zweier unendlich kleinen Größen A, B der Ordnung α bzw. $\beta \geq \alpha$ ist eine unendlich kleine Größe der Ordnung $\gamma \geq \alpha$ oder Null.

4. Die *Subtraktion* ist stets (und immer eindeutig) ausführbar; denn die Definition

$$B + X = A$$

der Differenz ergibt

$$X = A - B = [a_\varrho - b_\varrho].$$

5. Die *Multiplikation* zweier gemischter Größen ist erklärt durch die Formel

$$A \cdot B = [a_0 b_\varrho + a_1 b_{\varrho-1} + \dots + a_\varrho b_0].$$

Diese Verknüpfung ist (offenbar) kommutativ und (wie leicht zu erkennen) assoziativ, sowie distributiv zur Addition.

Ferner existiert ein Modul (Einselement)

$$E = [1, 0, 0, 0, \dots],$$

und es gilt der folgende

Satz. Das Produkt einer unendlich kleinen Größe A der Ordnung α mit einer unendlich kleinen Größe B der Ordnung β ist eine unendlich kleine Größe der Ordnung $\alpha + \beta$.

Aus diesem Satz ergibt sich durch vollständige Induktion: Von einer unendlich kleinen Größe X der Ordnung ξ ist die ν -te Potenz ($\nu = 0, 1, 2, \dots$) eine unendlich kleine Größe der Ordnung $\nu\xi$.

6. Die Division ist nicht immer möglich, da bereits die Aufgabe

$$[x_0, x_1, x_2, \dots] \cdot [0, 1, 0, 0, \dots] = [1, 0, 0, 0, \dots]$$

die unerfüllbare Forderung $x_0 \cdot 0 = 1$ stellt. Immerhin jedoch gilt der folgende

Satz. Der Quotient einer unendlich kleinen Größe A der Ordnung α durch eine unendlich kleine Größe B der Ordnung $\beta \leq \alpha$ ist eindeutig bestimmt als eine unendlich kleine Größe der Ordnung $\alpha - \beta$.

7. Die Ordnung innerhalb unseres Systems von gemischten Größen wird hergestellt durch die folgende Definition

$A < B$, wenn $a_0 = b_0, a_1 = b_1, \dots, a_{r-1} = b_{r-1}$ und $a_r < b_r$ für ein passendes r der Reihe $0, 1, 2, \dots$

Es bilden alsdann die drei Möglichkeiten

$$A < B, \quad A = B, \quad B < A$$

eine vollständige Disjunktion. Und unser komplexes Zahlssystem ist übergegangen in ein „eindimensionales“ nicht-archimedisches Zahlssystem²⁾. Denn man hat bei

$$A_r = [0, 1, 0, 0, \dots] \quad \text{für } r = 1, 2, \dots, n$$

für alle natürlichen n die Beziehung

$$\sum_{r=1}^n A_r = [0, n, 0, 0, \dots] < [1, 0, 0, 0, \dots].$$

8. Wir betrachten jetzt das in unserem gemischten Größensystem enthaltene Teilsystem der realen Größen

$$A = [a, 0, 0, 0, \dots], \quad B = [b, 0, 0, 0, \dots].$$

²⁾ Nach liebenswürdiger Mitteilung des Herrn van der Waerden ist auch bei Zahlkörpern der Fall der nicht-archimedischen Anordnung bereits behandelt worden: Vgl. sein ausgezeichnetes Lehrbuch *Moderne Algebra*, I. Teil, Berlin 1930, S. 211 oben, und die daselbst angegebene Literatur.

In diesem gelten die folgenden Gesetze:

$$A = B \Leftrightarrow a = b, \quad A < B \Leftrightarrow a < b,$$

$$A \pm B = [a \pm b, 0, 0, 0, \dots],$$

$$A \cdot B = [a \cdot b, 0, 0, 0, \dots],$$

$$A : B = [a : b, 0, 0, 0, \dots].$$

Das Teilsystem der realen Größen ist also vollständig isomorph zu dem System der reellen Zahlen. Erstere können daher die letzteren in allen rationalen Rechnungen vertreten und auch deren Namen übernehmen; wir schreiben speziell

$$x = [x, 0, 0, 0, \dots].$$

§ 2.

Reihen von unendlich kleinen Größen.

1. Der Begriff der Konvergenz. Es seien die Größen $C^{(v)}$ ($v = 1, 2, 3, \dots$) unendlich klein von mindestens erster Ordnung.

Definition 4). Die unendliche Reihe $\sum_{v=1}^{\infty} C^{(v)}$ heißt *konvergent*, wenn in ihrer Partialsumme

$$\sum_{v=1}^n C^{(v)} = [0, s_1^{(n)}, s_2^{(n)}, s_3^{(n)}, \dots]$$

die ersten r Komponenten für alle $n \geq$ einem passenden N ihre definitiven Werte haben. In Formeln:

Zu r existiert N , so daß $s_q^{(n)} = s_q$ für $q = 1, 2, \dots, r$, wenn nur $n \geq N$.

Die *notwendige und hinreichende Konvergenzbedingung* besteht offenbar darin, daß in $C^{(v)}$ die Anzahl der AnfangsnulLEN oder (m. a. W.) die Ordnung des Unendlichkleinseins mit v über alle Grenzen wächst.

Daher ist jede Reihe unmittelbar als konvergent erkennbar, deren v -tes Glied $C^{(v)}$ eine unendlich kleine Größe von mindestens v -ter Ordnung ist.

Es sei noch ausdrücklich darauf hingewiesen, daß diese Definition der Konvergenz zur Summe

$$S = [0, s_1, s_2, s_3, \dots]$$

insofern in Übereinstimmung steht mit dem Begriff der endlichen Summe, als die Gleichheit besteht:

$$C^{(1)} + C^{(2)} + \dots + C^{(n)} + 0 + 0 + 0 + \dots = \sum_{v=1}^n C^{(v)}.$$

4) Was die Konvergenzdefinition angeht, so ist dieselbe nach liebenswürdiger Mitteilung des Herrn van der Waerden enthalten in seinem in Fußnote 2) bereits zitierten Lehrbuch, nämlich auf S. 218 in Formel (1), wenn man für ϵ unendlich kleine Größen einsetzt.

2. *Differentialpotenzreihen*. Nunmehr seien die $a^{(\nu)}$ ($\nu = 1, 2, 3, \dots$) reale Größen (vgl. § 1, Nr. 8) und

$$dx = [0, x_1, x_2, x_3, \dots]$$

eine beliebige unendlich kleine Größe von mindestens erster Ordnung (die Bezeichnung ist mit Rücksicht auf spätere Bedürfnisse gewählt). Dann verstehen wir unter einer Differentialpotenzreihe eine unendliche Reihe der Gestalt

$$\sum_{\nu=1}^{\infty} a^{(\nu)} dx^{\nu}.$$

Die Konvergenz dieser Reihen ist ohne weiteres ersichtlich, da das allgemeine Glied nach § 1, Nr. 5, Schlußabsatz eine unendlich kleine Größe von mindestens ν -ter Ordnung ist.

3. Der Anschaulichkeit halber wollen wir die *Anfangskomponenten* einer solchen Reihe ausrechnen. Wir finden zunächst nach der Produktdefinition $dx = [0, x_1, x_2, x_3, \dots]$, $dx^2 = [0, 0, x_1^2, 2x_1x_2, \dots]$, $dx^3 = [0, 0, 0, x_1^3, \dots]$ und daher für die Reihensumme den Wert:

$$[0, a^{(1)}x_1, a^{(1)}x_2 + a^{(2)}x_1^2, a^{(1)}x_3 + a^{(2)} \cdot 2x_1x_2 + a^{(3)}x_1^3, \dots].$$

4. Weiter benötigen wir noch die Gestalt der *allgemeinen Komponente* der Reihensumme. Es ist zunächst

$$dx^{\nu} = [0^{(0)}, 0^{(1)}, \dots, 0^{(\nu-1)}, x_1^{\nu}, \varphi_{\nu+1}^{(\nu)}(x_0), \varphi_{\nu+2}^{(\nu)}(x_0), \dots],$$

wo die $\varphi_{\nu}^{(\nu)}(x_0)$ Summen von Produkten aus je ν Faktoren x_0 bedeuten, was sich leicht durch vollständige Induktion bestätigen läßt:

$$\begin{aligned} dx^{\nu+1} &= dx^{\nu} \cdot dx \\ &= [0^{(0)} \cdot x_0 + 0^{(1)} \cdot x_{0-1} + \dots + 0^{(\nu-1)} \cdot x_{0-\nu+1} + x_1^{\nu} \cdot x_{0-\nu} + \\ &\quad + \varphi_{\nu+1}^{(\nu)} \cdot x_{0-\nu-1} + \dots + \varphi_{\nu-1}^{(\nu)} \cdot x_1 + \varphi_{\nu}^{(\nu)} \cdot 0]. \end{aligned}$$

Und in der eckigen Klammer steht für $\varrho \leq \nu$ Null, für $\varrho = \nu + 1$ aber $x_1^{\nu+1}$ und für $\varrho > \nu + 1$ eine Summe von Produkten aus je $\nu + 1$ Faktoren x_0 . W. z. b. w.

Für die allgemeine Gestalt der ϱ -ten Komponente s_{ϱ} der Reihensumme findet man jetzt leicht:

$$(0) \quad s_{\varrho} = a^{(1)}\varphi_{\varrho}^{(1)} + a^{(2)}\varphi_{\varrho}^{(2)} + \dots + a^{(\varrho-1)}\varphi_{\varrho}^{(\varrho-1)} + a^{(\varrho)}x_1^{\varrho}.$$

Hieraus geht hervor, daß s_{ϱ} für $\varrho \leq \nu - 1$ nur aus Gliedern der Dimension $\leq \nu - 1$ besteht, und für $\varrho = \nu$ die Form hat

$$a^{(\nu)}x_1^{\nu} + \text{Glieder der Dimension } \leq \nu - 1.$$

Wir behandeln jetzt das Rechnen mit Differentialpotenzreihen:

5. Die *Vervielfachung* mit einer realen Konstanten c kann unter dem Summenzeichen vorgenommen werden:

$$c \sum_{v=1}^{\infty} a^{(v)} dx^v = \sum_{v=1}^{\infty} (c a^{(v)}) dx^v.$$

Beweis. Die Anfangskomponenten der zu multiplizierenden Reihe ergeben sich aus $\sum_{v=1}^n a^{(v)} dx^v$, wenn nur $n \geq N$; also ergeben sich die Anfangskomponenten der mit c multiplizierten Reihe aus

$$c \sum_{v=1}^n a^{(v)} dx^v = \sum_{v=1}^n (c a^{(v)}) dx^v, \text{ wenn nur } n \geq N,$$

d. h. aus der Partialsumme der unter dem Summenzeichen multiplizierten Reihe.

6. Die *Addition (Subtraktion)* zweier Differentialpotenzreihen kann gliedweise vorgenommen werden; in Formeln:

$$\sum_{v=1}^{\infty} a^{(v)} dx^v \pm \sum_{v=1}^{\infty} b^{(v)} dx^v = \sum_{v=1}^{\infty} (a^{(v)} \pm b^{(v)}) dx^v.$$

Beweis. Die Anfangskomponenten der Aggregatreihen ergeben sich aus $\sum_{v=1}^n a^{(v)} dx^v$ bzw. $\sum_{v=1}^n b^{(v)} dx^v$, wenn nur $n \geq N_1$ bzw. $n \geq N_2$; also ergeben sich die Anfangskomponenten der Summe (Differenz) aus

$$\sum_{v=1}^n a^{(v)} dx^v \pm \sum_{v=1}^n b^{(v)} dx^v = \sum_{v=1}^n (a^{(v)} \pm b^{(v)}) dx^v, \text{ wenn nur } n \geq \max(N_1, N_2),$$

d. h. aus der Partialsumme der gliedweise addierten (subtrahierten) Reihe.

7. Die *Multiplikation* zweier Differentialpotenzreihen erfolgt nach der Cauchyschen Regel; in Formeln:

$$\sum_{v=1}^{\infty} a^{(v)} dx^v \cdot \sum_{v=1}^{\infty} b^{(v)} dx^v = \sum_{v=2}^{\infty} (a^{(1)} b^{(v-1)} + a^{(2)} b^{(v-2)} + \dots + a^{(v-1)} b^{(1)}) dx^v.$$

Beweis. Die Anfangskomponenten jeder Faktorreihe ergeben sich wie vorstehend; also ergeben sich die Anfangskomponenten des Produktes aus

$$\sum_{v=1}^n a^{(v)} dx^v \cdot \sum_{v=1}^n b^{(v)} dx^v = \sum_{v=2}^n (a^{(1)} b^{(v-1)} + a^{(2)} b^{(v-2)} + \dots + a^{(v-1)} b^{(1)}) dx^v$$

+ Glieder höherer Ordnung, wenn nur $n \geq \max(N_1, N_2)$

und daher schließlich aus der Partialsumme der Cauchyschen Produktreihe.

8. Ein *Nullitätssatz*. Wenn die Gleichheit besteht

$$\sum_{r=1}^{\infty} a^{(r)} dx^r = 0 \text{ für einen einzigen Wert von } dx \text{ mit } x_1 \neq 0,$$

so verschwinden alle Koeffizienten der Differentialpotenzreihe.

Beweis. Aus der Voraussetzung folgt zunächst nach Nr. 3 dieses Paragraphen

$$a^{(1)} x_1 = 0, \text{ also } a^{(1)} = 0,$$

und, wenn das Verschwinden der ersten $\varrho - 1$ Koeffizienten bewiesen ist, aus der allgemeinen Gestalt (0) der ϱ -ten Komponente der Reihensumme (nach Nr. 4)

$$s_{\varrho} = 0 \cdot \varphi_{\varrho}^{(1)} + 0 \cdot \varphi_{\varrho}^{(2)} + \dots + 0 \cdot \varphi_{\varrho}^{(\varrho-1)} + a^{(\varrho)} x_1^{\varrho} = 0, \text{ also } a^{(\varrho)} = 0.$$

9. Ein *Identitätssatz*. Wenn die Gleichheit besteht

$$\sum_{r=1}^{\infty} a^{(r)} dx^r = \sum_{r=1}^{\infty} b^{(r)} dx^r \text{ für einen einzigen Wert von } dx \text{ mit } x_1 \neq 0,$$

so stimmen die Koeffizienten gleicher Nummer überein.

Beweis. Klar.

§ 3.

Die Deckfunktion einer reellen Funktion.

1. Das *spezielle Differential des Arguments*. Indem wir die früheren Bezeichnungen ein klein wenig abändern, setzen wir:

$$X = [x, x_1, x_2, x_3, \dots], \quad x = [x, 0, 0, 0, \dots], \quad dx = [0, x_1, x_2, x_3, \dots];$$

Wir verstehen also speziell unter dx eine beliebige unendlich kleine Größe von mindestens erster Ordnung, und es gilt stets die Beziehung

$$X = x + dx.$$

2. Das *erste spezielle Differential der Funktion* F der gemischten Veränderlichen X benötigen wir zunächst ebenfalls nur an den realen Stellen x und erklären es daselbst durch die Formel

$$dF(x) = F(x + dx) - F(x)$$

als eine richtiggehende Differenz.

3. Wir gehen jetzt daran, bei einer gegebenen, beliebig oft differentürierbaren reellen Funktion $f(x)$ die zugehörige gemischte „*Deckfunktion*“ $F(X)$

der gemischten Veränderlichen X zu konstruieren. Der Grundgedanke, von dem wir dabei ausgehen, besteht darin, daß zunächst einmal die Beziehungen

$$F(x) = f(x), \quad F'(x) = f'(x), \quad F''(x) = f''(x), \quad \dots$$

gelten sollen, und daß ferner an die Stelle der Formel

$$dF(x) = f'(x) dx,$$

die in einer Differentialrechnung mit Differentialen von nur erster Ordnung⁵⁾ gilt, jetzt, wo es sich um die Rechnung mit Differentialen aller Ordnungen handelt, die Formel tritt:

$$dF(x) = \sum_{v=1}^{\infty} v!^{-1} f^{(v)}(x) dx^v,$$

in welcher rechts die „*Taylor'sche Differentialpotenzreihe*“ steht. Wir kommen auf diese Weise zu der Definition

$$(1) \quad F(X) = F(x + dx) = F(x) + dF(x) = f(x) + \sum_{v=1}^{\infty} v!^{-1} f^{(v)}(x) dx^v.$$

Die Anfangskomponenten von $F(X)$ lassen sich nach § 2, Nr. 3 leicht angeben. Wichtiger ist die allgemeine Gestalt der ϱ -ten Komponente, die sich nach § 2, Nr. 4, Formel (0) ergibt:

$$y_{\varrho} = 1!^{-1} f'(x) \varphi_{\varrho}^{(1)}(x_0) + 2!^{-1} f''(x) \varphi_{\varrho}^{(2)}(x_0) + \dots + \\ + (\varrho - 1)!^{-1} f^{(\varrho-1)}(x) \varphi_{\varrho}^{(\varrho-1)}(x_0) + \varrho!^{-1} f^{(\varrho)}(x) x_1^{\varrho} \quad (\varrho = 1, 2, 3, \dots).$$

Aus der Definition (1) der Deckfunktion ist für $dx = 0$ ersichtlich, daß in der Tat die Beziehung gilt:

$$(2) \quad F(x) = f(x).$$

4. Ein interessanter *Nullitätssatz* besteht für die Deckfunktionen; er lautet: Wenn bei einer gegebenen Deckfunktion $F(X)$ die Beziehung besteht

$$F(x) = 0 \text{ für alle realen } x,$$

so gilt identisch

$$F(X) = 0 \text{ für alle gemischten } X.$$

Beweis. Die Voraussetzung deckt sich wegen (2) mit der Gleichung

$$f(x) = 0 \text{ für alle reellen } x.$$

Daher gilt aber

$$f^{(v)}(x) = 0 \text{ für alle reellen } x,$$

und jetzt folgt aus (1) die Behauptung.

⁵⁾ Vgl. meine frühere Arbeit: Modell einer Differentialrechnung mit aktual unendlich kleinen Größen erster Ordnung. Diese Annalen, Bd. 118, S. 251.

5. Der zugehörige *Identitätssatz* lautet: Wenn für zwei gegebene Deckfunktionen $F(X)$, $G(X)$ die Beziehung besteht

$$F(x) = G(x) \text{ für alle realen } x,$$

so gilt identisch

$$F(X) = G(X) \text{ für alle gemischten } X.$$

Der Beweis liegt auf der Hand.

6. Aus vorstehenden Betrachtungen folgt leicht der

Satz. Eine notwendige und hinreichende Bedingung für die Identität zweier Deckfunktionen $F(X)$, $G(X)$ für alle gemischten X besteht in der Identität der zugehörigen gedeckten reellen Funktionen $f(x)$, $g(x)$ für alle reellen x .

Beweis. Als notwendig und hinreichend ergibt sich trivialerweise bzw. nach dem Identitätssatz die Bedingung

$$F(x) = G(x) \text{ für alle realen } x,$$

und diese deckt sich wegen (2) mit der folgenden:

$$f(x) = g(x) \text{ für alle reellen } x.$$

7. Nunmehr verstehen wir unter der ν -ten *Abgeleiteten* einer gemischten Funktion $F(X)$ die Deckfunktion der ν -ten Ableitung $f^{(\nu)}(x)$ der zu $F(X)$ gehörigen gedeckten reellen Funktion $f(x)$.

Durch diese Definition sind wegen (2) die noch fehlenden Beziehungen

$$F^{(\nu)}(x) = f^{(\nu)}(x), \quad \nu = 1, 2, 3, \dots$$

gesichert. Ihretwegen kann dann die Taylorsche Differentialpotenzreihe folgendermaßen geschrieben werden:

$$dF(x) = \sum_{\nu=1}^{\infty} \nu!^{-1} F^{(\nu)}(x) dx^{\nu}.$$

8. Für den übernächsten Paragraphen notieren wir noch den folgenden

Hilfssatz. Es ist

$$dF(x) = \varepsilon_1,$$

wenn wir allgemein mit ε , eine unendlich kleine Größe von mindestens ν -ter Ordnung bezeichnen.

Beweis. Die Taylorsche Differentialpotenzreihe (vgl. oben, Nr. 3) zeigt, daß die linke Seite die Summe einer unendlichen Reihe von unendlich kleinen Größen von je mindestens erster Ordnung ist, woraus die Behauptung folgt.

§ 4.

Der Kalkül der Ableitungen.

1. Die *Methode*. Es handelt sich hier um den Nachweis der Gleichheit zwischen einer gegebenen Abgeleiteten $F^{(v)}(X)$ und einem durch Rechnung gefundenen Ausdruck $G_v(X)$. Zu diesem Zweck gehen wir aus von der gegebenen Taylorschen Differentialpotenzreihe (§ 3, Nr. 7, am Schluß)

$$dF(x) = \sum_{v=1}^{\infty} v!^{-1} F^{(v)}(x) dx^v$$

und vergleichen dieselbe mit einer durch Rechnung gefundenen Entwicklung

$$dF(x) = \sum_{v=1}^{\infty} v!^{-1} G_v(x) dx^v.$$

Dann liefert der Identitätssatz für Differentialpotenzreihen (§ 2, Nr. 9) die Gleichheit

$$F^{(v)}(x) = G_v(x) \text{ für alle realen } x.$$

Und hieraus folgt nach dem Identitätssatz für Deckfunktionen (§ 3, Nr. 5) die Gleichheit

$$F^{(v)}(X) = G_v(X) \text{ für alle gemischten } X.$$

2. Für die *Summe (Differenz)* zweier Deckfunktionen sei

$$F(X) = U(X) \pm V(X).$$

Daher wird

$$\begin{aligned} dF(x) &= \{U(x+dx) \pm V(x+dx)\} - \{U(x) \pm V(x)\} \\ &= \{U(x+dx) - U(x)\} \pm \{V(x+dx) - V(x)\} \\ &= dU(x) \pm dV(x) = \sum_{v=1}^{\infty} v!^{-1} U^{(v)}(x) dx^v \pm \sum_{v=1}^{\infty} v!^{-1} V^{(v)}(x) dx^v \\ &= \sum_{v=1}^{\infty} v!^{-1} \{U^{(v)}(x) \pm V^{(v)}(x)\} dx^v, \end{aligned}$$

und das Resultat lautet

$$F^{(v)}(X) = U^{(v)}(X) \pm V^{(v)}(X).$$

3. Für das *Produkt* zweier Deckfunktionen sei

$$F(X) = U(X) \cdot V(X).$$

Dann wird nach üblicher Rechnung

$$\begin{aligned}
 dF(x) &= U(x+dx) \cdot V(x+dx) - U(x) \cdot V(x) \\
 &= U(x) dV(x) + dU(x) \cdot V(x) + V(x) dU(x) \\
 &= U(x) \cdot \sum_{v=1}^{\infty} v!^{-1} V^{(v)}(x) dx^v + \sum_{v=1}^{\infty} v!^{-1} U^{(v)}(x) dx^v \cdot \sum_{v=1}^{\infty} v!^{-1} V^{(v)}(x) dx^v + \\
 &\quad + V(x) \cdot \sum_{v=1}^{\infty} v!^{-1} U^{(v)}(x) dx^v \\
 &= \sum_{v=1}^{\infty} v!^{-1} \left\{ U(x) V^{(v)}(x) + \binom{v}{1} U'(x) V^{(v-1)}(x) + \dots + \right. \\
 &\quad \left. + \binom{v}{v-1} U^{(v-1)}(x) V'(x) + U^{(v)}(x) V(x) \right\} dx^v
 \end{aligned}$$

und das Resultat lautet

$$\begin{aligned}
 F^{(v)}(X) &= U(X) V^{(v)}(X) + \binom{v}{1} U'(X) V^{(v-1)}(X) + \dots + \\
 &\quad + \binom{v}{v-1} U^{(v-1)}(X) V'(X) + U^{(v)}(X) V(X).
 \end{aligned}$$

4. Um den *Quotienten* behandeln zu können, genügt es, die *Reziproke* zu erledigen: Sei also

$$F(X) = \frac{1}{V(X)}.$$

Daraus folgt

$$\begin{aligned}
 dF(x) &= \frac{1}{V(x+dx)} - \frac{1}{V(x)} = -\frac{V(x+dx) - V(x)}{V(x)V(x+dx)} = \frac{-dV(x)}{V(x)^2 \left(1 + \frac{dV(x)}{V(x)}\right)} \\
 &= \sum_{\mu=1}^{\infty} \frac{(-1)^{\mu}}{V(x)^{\mu+1}} (dV(x))^{\mu} = \sum_{\mu=1}^{\infty} \frac{(-1)^{\mu}}{V(x)^{\mu+1}} \left(\sum_{v=1}^{\infty} v!^{-1} V^{(v)}(x) dx^v \right)^{\mu}.
 \end{aligned}$$

Nach dem Zusammenfassen gleich hoher Potenzen von dx ist wieder die Koeffizientenvergleichung möglich, deren Anfangsresultate (mit X statt x) lauten:

$$F'(X) = \frac{-1}{V(X)^2} \cdot V'(X), \quad F''(X) = \frac{-1}{V(X)^3} \cdot V''(X) + \frac{2}{V(X)^3} \cdot V'(X)^2, \dots$$

5. Bei der *Funktion von einer Funktion* ist

$$F(X) = \Phi(U(X)).$$

Dann wird

$$dF(x) = \Phi(U(x+dx)) - \Phi(U(x)) = \Phi(U(x) + dU(x)) - \Phi(U(x)).$$

Hierin ist $dU(x)$ eine unendlich kleine Größe mindestens erster Ordnung, also von derselben Beschaffenheit wie dx , und $U(x)$ ist eine reale Größe, so daß wir erhalten

$$\begin{aligned} dF(x) &= \sum_{\mu=1}^{\infty} \mu!^{-1} \Phi^{(\mu)}(U(x)) \cdot (dU(x))^{\mu} \\ &= \sum_{\mu=1}^{\infty} \mu!^{-1} \Phi^{(\mu)}(U(x)) \cdot \left(\sum_{\nu=1}^{\infty} \nu!^{-1} U^{(\nu)}(x) dx^{\nu} \right)^{\mu}. \end{aligned}$$

Wiederum ist nach erfolgter Ordnung⁶⁾ nach Potenzen von dx die Koeffizientenvergleichung möglich, deren Anfangsergebnisse lauten:

$$F'(X) = \Phi'(U(X)) \cdot U'(X),$$

$$F''(X) = \Phi'(U(X)) \cdot U''(X) + \Phi''(U(X)) \cdot U'(X)^2, \dots$$

6. Bei der umgekehrten Funktion ist

$$Y = F(X), \quad X = \Phi(Y); \quad \text{also } X = \Phi(F(X)).$$

Setzt man darin $X = x + dx$, so wird (unter Benutzung von Nr. 6)

$$\begin{aligned} x + dx &= \Phi(F(x + dx)) = \Phi(F(x) + dF(x)) \\ &= \Phi(F(x)) + \sum_{\mu=1}^{\infty} \mu!^{-1} \Phi^{(\mu)}(F(x)) \cdot \left(\sum_{\nu=1}^{\infty} \nu!^{-1} F^{(\nu)}(x) dx^{\nu} \right)^{\mu}. \end{aligned}$$

Nach Ordnung und Koeffizientenvergleichung ergibt sich (zunächst für kleine x , und daher auch für große X):

$$X = \Phi(F(X)),$$

$$1 = \Phi'(F(X)) \cdot F'(X),$$

$$0 = \Phi'(F(X)) \cdot F''(X) + \Phi''(F(X)) \cdot F'(X)^2,$$

$$\dots \dots \dots$$

Aus diesem System ergibt sich zunächst $F(X)$ und weiterhin $F'(X), F''(X), \dots$

⁶⁾ Aus Gründen der Vollständigkeit müssen wir hier eine reihentheoretische Bemerkung einfügen.

Man erkennt zunächst, daß die beiden vorstehenden (miteinander übereinstimmenden) Reihen aus den davor angegebenen Gründen in der vorliegenden Anordnung konvergent sind.

Weiter ergibt sich dann, daß die Umordnung der Zeilensumme in die Spaltensumme erlaubt ist, also daß

$$\sum_{\mu=1}^{\infty} \left(\sum_{\nu=\mu}^{\infty} c^{(\mu, \nu)} dx^{\nu} \right) = \sum_{\nu=1}^{\infty} \left(\sum_{\mu=1}^{\nu} c^{(\mu, \nu)} dx^{\nu} \right).$$

Denn für $\varrho = 1, 2, \dots, r$ ergeben sich die zugehörigen Komponenten der beiden Reihensummen aus

$$\sum_{\mu=1}^r \left(\sum_{\nu=\mu}^r c^{(\mu, \nu)} dx^{\nu} \right) \quad \text{bzw.} \quad \sum_{\nu=1}^r \left(\sum_{\mu=1}^{\nu} c^{(\mu, \nu)} dx^{\nu} \right),$$

also aus einer und derselben gemischten Zahl. Und jede Komponente behält ihren Wert bei, wenn r vergrößert wird.

§ 5.

Der Kalkül der Differentialquotienten.

Nachdem wir vorstehend den Kalkül der Ableitungen behandelt haben, können wir jetzt darangehen, Formeln für die ν -ten Differentialquotienten der zusammengesetzten Funktionen abzuleiten, die allerdings (im Gegensatz zu den für die Ableitungen erhaltenen Formeln) immer nur bis auf unendlich kleine Größen von mindestens erster Ordnung gelten.

Neben den Differentialen des § 3 benötigen wir jetzt noch die folgenden:

1. Das *allgemeine Differential des Arguments*

$$X = [x, x_1, x_2, x_3, \dots] \text{ heißt } dX = [0, x'_1, x'_2, x'_3, \dots];$$

es ist (wie seinerzeit dx) eine beliebige unendlich kleine Größe von mindestens erster Ordnung.

2. Das *allgemeine ν -te Differential der Funktion* der gemischten Variablen X an der Stelle X ist erklärt durch die Formeln

$$dF(X) = F(X + dX) - F(X); \quad d^\nu F(X) = d(d^{\nu-1} F(X)), \quad \nu = 2, 3, 4, \dots$$

Es ist also eine ganz gewöhnliche ν -te Differenz

$$d^\nu F(X) = \sum_{\alpha=0}^{\nu} (-1)^{\nu-\alpha} \binom{\nu}{\alpha} F(X + \alpha dX), \quad \nu = 1, 2, 3, \dots$$

3. Nunmehr ergibt sich (in der Bezeichnung von § 3, Nr. 8) der folgende Satz 1. *Es ist*

$$d^\nu F(X) = F^{(\nu)}(X) \cdot dX^\nu + \varepsilon_{\nu+1}.$$

Beweis. Wir führen zunächst eine Formel an, die wir früher abgeleitet haben und jetzt zum Beweis nötig haben: Nach § 2, Nr. 4 gilt bei

$$dx = [0, x_1, x_2, x_3, \dots]$$

zunächst

$$dx^\nu = [0^{(\nu)}, 0^{(1)}, \dots, 0^{(\nu-1)}, x_1^{(\nu)}, \varphi_{\nu+1}^{(\nu)}(x_0), \varphi_{\nu+2}^{(\nu)}(x_0), \dots]$$

— wo die $\varphi_\nu^{(\nu)}(x_0)$ Summen von Produkten aus je ν Faktoren x_u bedeuten —, so daß sich weiter ergibt

$$dX^\nu = [0^{(\nu)}, 0^{(1)}, \dots, 0^{(\nu-1)}, (x'_1)^\nu, \varphi_{\nu+1}^{(\nu)}(x'_0), \varphi_{\nu+2}^{(\nu)}(x'_0), \dots].$$

Nunmehr gehen wir zur Bildung des ν -ten Differentials über: Man hat zunächst

$$dx + \alpha dX = [0, x_1 + \alpha x'_1, x_2 + \alpha x'_2, x_3 + \alpha x'_3, \dots]$$

und daher

$$F(X + \alpha dX) = F(x + (dx + \alpha dX)) = [F(x), y_1, y_2, y_3, \dots],$$

wo nach § 3, Nr. 3 und Nr. 7

$$y_\varrho = 1!^{-1} F^{(\varrho)}(x) \varphi_\varrho^{(1)}(x_0 + \alpha x'_0) + \dots + (\varrho - 1)!^{-1} F^{(\varrho-1)}(x) \varphi_\varrho^{(\varrho-1)}(x_0 + \alpha x'_0) \\ + \varrho!^{-1} F^{(\varrho)}(x) \cdot (x_1 + \alpha x'_1)^\varrho.$$

Darin ist $F(x)$ und (nach dem Schlußsatz von § 2, Nr. 4) y_ϱ für $\varrho = 1, 2, \dots, \nu - 1$ in α vom Grade $\leq \nu - 1$, so daß die zugehörigen Komponenten bei der Bildung der ν -ten Differenz verschwinden. Weiter ist für $\varrho = \nu$

$$y_\nu = \nu!^{-1} F^{(\nu)}(x) \alpha^\nu (x'_1)^\nu + \text{Glieder, die in } \alpha \text{ vom Grade } \leq \nu - 1,$$

so daß als ν -te Komponente von $d^\nu F(X)$ lediglich verbleibt

$$\nu!^{-1} F^{(\nu)}(x) \nu! (x'_1)^\nu.$$

Mit Rücksicht auf den oben berechneten Wert von dX^ν erhalten wir daher schließlich

$$d^\nu F(X) = F^{(\nu)}(x) \cdot dX^\nu + \varepsilon_{\nu+1} = (F^{(\nu)}(X) - dF^{(\nu)}(x)) \cdot dX^\nu + \varepsilon_{\nu+1},$$

und dieser Ausdruck ist wegen des Hilfssatzes in § 3, Nr. 8

$$= F^{(\nu)}(X) \cdot dX^\nu - \varepsilon_1 \cdot dX^\nu + \varepsilon_{\nu+1} = F^{(\nu)}(X) \cdot dX^\nu + \varepsilon_{\nu+1},$$

also gleich dem behaupteten Wert.

4. Leicht folgt jetzt der

Satz 2. *Es ist, wenn jetzt noch obendrein $x'_1 \neq 0$, also m. a. W. dX unendlich klein von genau erster Ordnung,*

$$\frac{d^\nu F(X)}{dX^\nu} = F^{(\nu)}(X) + \varepsilon_1.$$

Beweis. Nach dem Satz der Nr. 6 des § 1 ist in der Behauptung des vorigen Satzes

$$\frac{\varepsilon_{\nu+1}}{dX^\nu} = \varepsilon_1, \quad \varepsilon_{\nu+1} = \varepsilon_1 \cdot dX^\nu, \quad \text{also} \quad d^\nu F(X) = [F^{(\nu)}(X) + \varepsilon_1] \cdot dX^\nu,$$

woraus die Behauptung von Satz 2 folgt.

Es genügt nunmehr, zwei Beispiele vorzuführen, um die allgemeine Beweismethode zu erkennen, die auf dem Satz 2 beruht:

5. Zunächst das Produkt

$$F(X) = U(X) \cdot V(X).$$

Hier ergibt sich

$$\frac{d^\nu F(X)}{dX^\nu} = F^{(\nu)}(X) + \varepsilon_1 = \sum_{\alpha=0}^{\nu} \binom{\nu}{\alpha} U^{(\alpha)}(X) \cdot V^{(\nu-\alpha)}(X) + \varepsilon_1,$$

und hieraus folgt bei nochmaliger Anwendung von Satz 2

$$\frac{d^\nu F(X)}{dX^\nu} = \sum_{\alpha=0}^{\nu} \binom{\nu}{\alpha} \frac{d^\alpha U(X)}{dX^\alpha} \cdot \frac{d^{\nu-\alpha} V(X)}{dX^{\nu-\alpha}} + \varepsilon_1.$$

6. Schließlich betrachten wir die *Funktion von einer Funktion* und speziell die früher gefundene Formel für $F''(X)$; aus ihr folgt bei

$$F(X) = \Phi(U(X))$$

die Gleichungskette

$$\begin{aligned}\frac{d^2 F(X)}{dX^2} &= F''(X) + \varepsilon_1 \\ &= \Phi'(U(X)) \cdot U''(X) + \Phi''(U(X)) \cdot U'(X)^2 + \varepsilon_1 \\ &= \Phi'(U(X)) \cdot \frac{d^2 U(X)}{dX^2} + \Phi''(U(X)) \cdot \left(\frac{dU(X)}{dX}\right)^2 + \varepsilon_1,\end{aligned}$$

womit auch hier die gewünschte Formel gefunden ist.

§ 6.

Schlußwort.

Zum Schlusse der vorliegenden Arbeit machen wir den Leser darauf aufmerksam, daß es eine interessante Fragestellung ist, den Vorrat an speziellen Funktionen zu betrachten, der sich unter unseren Deckfunktionen vorfindet. Wir beschränken uns hier auf die Vorführung zweier Beispiele:

1. Die *Exponentialfunktion*. Sie ist die Deckfunktion $F(X)$ der reellen Funktion $f(x) = e^x$ mit den Ableitungen $f^{(v)}(x) = e^x$, so daß sich aus § 3, Nr. 3, Gleichung (1) ergibt

$$e^X = F(X) = f(x) + \sum_{v=1}^{\infty} v!^{-1} f^{(v)}(x) dx^v = e^x \cdot \left(1 + \sum_{v=1}^{\infty} v!^{-1} dx^v\right).$$

Für diese Funktion gilt das Additionstheorem, das z. B. folgendermaßen durch Rechnung bewiesen werden kann:

$$\begin{aligned}e^{X'+X''} &= e^{x'+x''} \cdot \left(1 + \sum_{v=1}^{\infty} v!^{-1} (dx' + dx'')^v\right) \\ &= e^{x'+x''} \cdot \left(1 + \sum_{v=1}^{\infty} \sum_{\alpha=0}^v \alpha!^{-1} (dx')^\alpha \cdot (v-\alpha)!^{-1} (dx'')^{v-\alpha}\right) \\ &= e^{x'} \cdot \left(1 + \sum_{v=1}^{\infty} v!^{-1} (dx')^v\right) \cdot e^{x''} \cdot \left(1 + \sum_{v=1}^{\infty} v!^{-1} (dx'')^v\right) = e^{x'} \cdot e^{x''}.\end{aligned}$$

2. Der *natürliche Logarithmus*. Er ist die Deckfunktion $F(X)$ der reellen Funktion $f(x) = \log x$ ($x > 0$) mit den Ableitungen $f^{(v)}(x) = (-1)^{v-1} (v-1)! x^{-v}$, so daß sich nach der vorgenannten Formel ergibt

$$\text{Log } X = F(X) = \log x + \sum_{v=1}^{\infty} \frac{(-1)^{v-1}}{v} \cdot \left(\frac{dx}{x}\right)^v.$$

Ebenso gilt

$$\text{Log } X' + \text{Log } X'' = \text{Log } (X'X'').$$

(Eingegangen am 3. 5. 1941.)

Zur Theorie der elliptischen Differentialgleichungen. II.

Von

Georg Tautz in Breslau.

Inhaltsübersicht.

	Seite
Einleitung	733
1. Das klassische Problem	735
§ 1. Der Operator $L_{f,x}$	735
§ 2. Konstruktion von Lösungen bezüglich Ω	738
§ 3. Verhalten auf dem Rande bei kleinen Gebieten	741
§ 4. Verhalten auf dem Rande bei beliebigen Gebieten	744
§ 5. Bedingung VI. Klassische Lösung der Randwertaufgabe	745
2. Das verallgemeinerte Problem	750
§ 6. Normierte Potentiale	750
§ 7. Greensche Funktion. Kapazität	752
§ 8. Vergleichung mit der Kapazität bei Δu	754
§ 9. Abschluß der allgemeinen Theorie. Der Satz über die regulären Randpunkte	757
3. Anwendung auf elliptische Differentialgleichungen	760
§ 10. Die Bedingungen I—VII	760
§ 11. Ein Eindeutigkeitssatz	761
§ 12. Zwei weitere Hilfssätze	763
§ 13. Abschluß des Beweises	769

Einleitung.

Im ersten Teil der vorliegenden Abhandlung¹⁾ wurde im wesentlichen das Dirichletsche Problem bei der Funktionalgleichung

$$L(u) = - \int_s \frac{\partial u}{\partial n} ds + \int_{cs} \left(\frac{\partial u}{\partial x} dA + \frac{\partial u}{\partial y} dB + u dC \right)$$

für hinreichend kleine und hinreichend reguläre Gebiete gelöst.

¹⁾ Vgl. Math. Annalen 117, S. 694—727, im folgenden mit I zitiert.

Im folgenden wird das Dirichletsche Problem in seiner allgemeinsten Form²⁾ für beschränkte Gebiete behandelt. Die Ausführungen sind in drei Abschnitte gegliedert. In den beiden ersten wird das Problem zunächst unter einem wesentlich allgemeineren Gesichtspunkt, nämlich ohne Bezugnahme auf eine bestimmte Funktionalgleichung in Angriff genommen. Anlaß und Ausgangspunkt hierfür ist die bekannte Arbeit von O. Perron „Eine neue Behandlung der ersten Randwertaufgabe für $\Delta u = 0$ “³⁾ die seitdem bereits mehrfach Gegenstand interessanter Untersuchungen geworden ist.

Für ein System von Kreisen κ in dem beschränkten Gebiet Ω_0 ist eine lineare Operation L_{κ} definiert, d. h. eine gewisse stetige Funktion, die auf dem Rande des Kreises die stetigen Randwerte f annimmt. Der Operator L ist im wesentlichen zwei qualitativen Bedingungen unterworfen. Er erfüllt ein bestimmtes Maximumprinzip und besitzt eine gewisse Kompaktheitseigenschaft (vgl. die Eigenschaften I bis V in § 1).

Unter diesen allgemeinen Voraussetzungen gelingt es bereits durch Ausgestaltung des Perronschen Existenzbeweises, den Definitionsbereich der Operation auf beliebige Gebiete zu erweitern und jeder beschränkten Randfunktion (mindestens) eine solche Funktion im Innern zuzuordnen. In § 5 wird eine neue Bedingung eingeführt, mittels deren für den Operator L die Lösung des Dirichletschen Problems für Gebiete mit stetig gekrümmten Randkomponenten im klassischen Sinne gesichert werden kann³⁾.

Dem zweiten Abschnitt, welcher den von N. Wiener und O. D. Kellogg eingeführten Begriff der Kapazität zur Charakterisierung des Randes benutzt, liegen zwei weitere wesentlich schärfere Bedingungen zugrunde (vgl. § 7). Hervorgehoben sei das Ergebnis, daß man unter diesen immer noch sehr allgemein gehaltenen Bedingungen bereits vollständig die bekannten Resultate

²⁾ Vgl. hierzu am besten O. D. Kellogg, Foundations of potential theory.

³⁾ Math. Zeitschr. 18. Unter den einschlägigen Arbeiten seien hervorgehoben: C. Carathéodory, On Dirichlets Problem. Amer. Journ. of Math. 59 (1937); M. Brelot, Familles de Perron. Acta Szeged IX (1938/40); G. Fubini, Sopra una nuova classe di problemi al contorno. Rend. Pal. 61 (1937). Die vorliegende Verallgemeinerung der Perronschen Arbeit ist die Ausgestaltung eines Vortrages, den ich 1933 im Mathematischen Seminar zu Breslau gehalten habe. Insofern es sich auch hier um die Befreiung von einer speziellen Funktionalgleichung handelt, berührt sie sich mit der soeben zitierten Arbeit von Fubini. Bei dieser handelt es sich jedoch nicht um eine Verallgemeinerung des Poissonschen Integrals selbst, sondern nur der Mittelwerteigenschaft desselben. Insofern erscheint der Ansatz von Fubini noch allgemeiner. Er hat dafür allerdings den Nachteil, daß er an ein festes Gebiet geknüpft ist, für das die Randwertaufgabe zu lösen ist; außerdem ist das Ergebnis insofern noch nicht befriedigend, als dort die Erfüllung der Randbedingung nur unter der Bedingung bewiesen wird, daß sie auch von einer gewissen Ober- und Unterfunktion erfüllt wird, während hier die wirkliche Existenz allgemeinerer Gebiete bewiesen wird, für welche das Problem sicher lösbar ist. Methodisch werden wir uns mehr an die zitierte Arbeit von Brelot anlehnen.

der Potentialtheorie erhält, insbesondere also die Identität der regulären Randpunkte für den Operator L mit denen für Δu .

Im dritten Abschnitt wird unter Zugrundelegung der Bedingungen $(C_{a,b})$ von I (S 714) gezeigt, daß der dort eingeführte verallgemeinerte elliptische Operator $L(u)$ alle aufgestellten Forderungen erfüllt, und mithin für ihn das Dirichletsche Problem in demselben Umfange lösbar ist, wie in der Potentialtheorie. Die Befreiung von der schärferen Bedingung (C_s) wird gelegentlich einer Fortführung dieser Untersuchungen erfolgen.

1. Das klassische Problem.

§ 1.

Der Operator $L_{f,x}$.

In einem gewissen Grundgebiet Ω_0 (offenes Kontinuum) sei ein System \mathfrak{S} von Kreisen κ wie folgt definiert:

Jeder Kreis κ liegt mit seinem Rande in Ω_0 . Gehört κ zu \mathfrak{S} , so gehört auch jeder Kreis in κ zu \mathfrak{S} . Jeder Punkt von Ω ist Mittelpunkt wenigstens eines Kreises aus \mathfrak{S} .

Sei κ irgendein Kreis aus \mathfrak{S} . Jeder stetigen Funktion f der Bogenlänge auf dem Rande von κ wird im Innern eine Funktion

$$L_{f,\kappa}(P)$$

zugeordnet, die folgende Eigenschaften besitzt:

I. Stetigkeitsbedingung. $L_{f,\kappa}(P)$ ist stetig in κ und schließt sich stetig an die Randwerte f an.

II. Eindeutigkeitsbedingung. Liegt κ' in κ , und setzt man $L_{f,\kappa}(P) = u(P)$, so ist in κ'

$$L_{u,\kappa'}(P) = L_{f,\kappa}(P).$$

III. Linearität. Bei festem κ gilt für beliebige reelle Zahlen λ, μ

$$L_{\lambda f + \mu g, \kappa} = \lambda L_{f,\kappa} + \mu L_{g,\kappa}.$$

IV. Maximumprinzip. Sei $L_{f,\kappa}(P) = u(P)$. Ist in einem inneren Punkte P_0 von κ , bzw. in einer Vollumgebung U von P_0

$$u(P_0) > 0, \quad u(P) \leq u(P_0) \quad (P \in U),$$

so ist in U

$$u(P) \equiv u(P_0).$$

V. Kompaktheitsbedingung. κ sei ein fester Kreis aus \mathfrak{S} . Dann ist die Gesamtheit der Funktionen $L_{f,\kappa}$, deren Randfunktionen *gleichmäßig beschränkt* sind, in jedem inneren Punkte *gleichgradig stetig*.

Wir nennen eine Funktion $u(P)$, die für jeden Kreis aus \mathfrak{S} der Gleichung $L_{u,\kappa}(P) = u(P)$ genügt, eine Lösung in Ω_0 oder kurz Lösung.

1. Folgerungen aus Bedingung IV. Es gilt ein zu IV analoges *Minimumprinzip*: Nimmt eine Lösung in P ein negatives Minimum an, so ist sie in einer Umgebung von P konstant.

Es gilt

$$\max_{P \in \kappa} |L_{f,\kappa}(P)| = \max |f|$$

mit der Verschärfung bei nichtkonstantem f für innere Punkte von κ

$$|L_{f,\kappa}(P)| < \max |f|.$$

Konvergieren die stetigen Funktionen f_n gleichmäßig gegen f , so konvergiert auch $L_{f_n,\kappa}(P)$ gleichmäßig gegen $L_{f,\kappa}(P)$.

Beim Beweis der ersten Aussage ist auch III zu benutzen.

Definition 1. Ω sei eine beliebige beschränkte offene Punktmenge in Ω_0 . Eine in Ω stetige Funktion $v(P)$ heißt *L-konkav* (bzw. *L-konvex*) in Ω , wenn für jeden Kreis κ aus \mathfrak{S} in Ω die Ungleichung

$$v(P) \geq L_{v,\kappa}(P) \quad (v(P) \leq L_{v,\kappa}(P)) \quad (P \in \kappa)$$

gilt.

Durch Multiplikation mit -1 gehen *L-konkave* und *L-konvexe* Funktionen ineinander über. Für das folgende ist der Umstand von Wichtigkeit, daß es für die *L-Konkavität* einer Funktion v bereits genügt, daß es um jeden Punkt P von Ω als Mittelpunkt überhaupt einen Kreis κ_P in Ω gibt derart, daß die obige Ungleichung für κ_P und jeden kleineren konzentrischen Kreis gilt. Andernfalls müßte nämlich für wenigstens einen Kreis κ in Ω und einen Punkt P in κ gelten $v(P) < L_{v,\kappa}(P)$. Die Funktion $V = v - L_{v,\kappa}$ hätte dann in einem Punkte P_1 von κ ein negatives Minimum. P_1 sei so gewählt, daß in jeder Nähe Punkte liegen, für welche $V(P) > V(P_1)$ ist. κ^* sei nun ein hinreichend kleiner Kreis um P_1 in κ , für welchen $v(P_1) \geq L_{v,\kappa^*}(P_1)$ gilt. Dann ist, $-L_{v,\kappa}(P) = w$ gesetzt,

$$L_{V,\kappa^*} = L_{v,\kappa^*} + L_{w,\kappa^*} = L_{v,\kappa^*} + w \quad (\text{wegen II}).$$

Also

$$V(P_1) = v(P_1) + w(P_1) \geq L_{v,\kappa^*}(P_1) + w(P_1) = L_{V,\kappa^*}(P_1).$$

Andererseits ist $V(P_1)$ das Minimum und, da V auf dem Rande von κ^* nicht konstant ist, wegen IV $L_{V,\kappa^*}(P_1) > V(P_1)$ im Widerspruch zur letzten Ungleichung.

Definition 2. Als obere (bzw. untere) *Hülle* $H(P)$ von Funktionen $w(P)$ über einer Punktmenge w bezeichnen wir die Funktion, welche in jedem Punkte gleich der oberen (bzw. unteren) Grenze der $w(P)$ ist.

Die meisten der folgenden kleinen Sätze sind nur für L -konvexe Funktionen formuliert, doch gilt natürlich entsprechendes auch für L -konkave.

2. Die Randfunktionen f seien gleichmäßig beschränkt. Dann ist die obere Hülle H von $L_{f,x}$ in x L -konvex.

Beweis. Wegen der gleichgradigen Stetigkeit der $L_{f,x}$ nach V ist H im Innern von x jedenfalls stetig. Liegt x^* im Innern von x , so können wir also L_{H,x^*} bilden, und es ist wegen IV, wenn $L_{f,x}(P) = v(P)$ gesetzt wird,

$$L_{H,x^*}(P) \geq L_{v,x^*}(P) = v(P).$$

Da dies für jedes $v(P)$ gilt, so auch für ihre obere Hülle; mithin ist

$$L_{H,x^*}(P) \geq H(P),$$

w. z. b. w.

Ganz ähnlich beweist man

3. Die obere Hülle H von in Ω L -konvexen Funktionen ist, falls in Ω stetig, wiederum L -konvex.

4. Ist v L -konvex und in jedem Randpunkte p von Ω der obere limes $\overline{v(p)}$ nicht positiv, so ist in Ω überall

$$v(P) \leq 0.$$

Beweis wie bei dem entsprechenden Satze bei subharmonischen Funktionen.

5. Ist $v(P)$ L -konvex, und ersetzt man in einem Kreise x aus \mathfrak{S} , der auch in Ω liegt, $v(P)$ durch $L_{v,x}(P)$, so ist die hierdurch entstehende Funktion $v_x(P)$ wiederum L -konvex in Ω .

Dies beweist man unter sinngemäßer Übertragung und leichter Abänderung genau wie den Hilfssatz 2 bei Perron. Man hat dabei noch folgendes zu benutzen. Sei $u = L_{f,x}$ auf dem Rande s eines in x gelegenen Gebietes ω nicht negativ

$$u(P) \geq 0 \quad (P \in s),$$

dann ist auch überall in ω

$$u(P) \geq 0.$$

Andernfalls hätte u in ω ein negatives Minimum μ , und man folgert leicht aus dem Minimumprinzip, daß die Punktmenge, auf welcher

$$u(P) = \mu$$

ist, mindestens einen Häufungspunkt auf s hat, was der Voraussetzung $u \geq 0$ auf s widerspräche.

§ 2.

Konstruktion von Lösungen bezüglich Ω .

Es sei jetzt auf dem Rande S der offenen beschränkten Punktmenge Ω in Ω_n eine beschränkte Funktion $f(p)$ ⁴⁾ vorgegeben.

$$(1, 2) \quad m \leq f \leq M.$$

Definition 3. Wir nennen mit Perron⁵⁾ eine in Ω definierte stetige Funktion $\varphi(P)$ Oberfunktion (kurz O.F.), wenn sie in Ω L -konkav ist und für jeden Randpunkt p ihr unterer limes $\underline{\varphi}(p)$ von $f(p)$ nicht übertroffen wird. Entsprechend heißt $\varphi(P)$ Unterfunktion (U.F.), wenn sie in Ω L -konvex ist, und wenn

$$(2, 2) \quad \overline{\varphi}(p) \leq f(p).$$

Dann gelten folgende Sätze:

6. Jede nichtpositive Konstante λ , die m nicht übertrifft, ist U.F. Zu zeigen ist nur

$$\lambda \leq L_{i,x}.$$

Dies folgt aber sofort aus IV.

7.

$$(3, 2) \quad \varphi \leq \text{Max}(0, M) = \mu.$$

$\mu - \varphi$ ist offenbar L -konkav; wegen (2, 2) hat die Funktion in jedem Randpunkt einen nichtnegativen unteren limes und ist mithin nach 4. überall nichtnegativ.

Wir nehmen in der Folge an

$$(4, 2) \quad m \leq 0 \leq M.$$

8. $\text{Max}(m, \varphi)$ ist U.F.

Die Bedingung am Rande ist offenbar erfüllt. Außerdem ist die Funktion stetig, also nach 3. wieder konvex.

Es sei jetzt $\underline{H}(P)$ die obere Hülle der Unterfunktionen. Wegen 7. und (4) ist

$$(5, 2) \quad \underline{H}(P) \leq M.$$

Ersetzen wir alle U.F. $\varphi(P)$ durch

$$\varphi_m(P) = \text{Max}(\varphi, m),$$

⁴⁾ Mit kleinen lateinischen Buchstaben bezeichnen wir fortan Randpunkte.

⁵⁾ Die oben gegebene Definition der Oberfunktionen stimmt mit der von Brelot benutzten überein, weicht also insofern von der Perronschen ab, als die Funktionen auf dem Rande nicht stetig zu sein brauchen.

so ist offenbar deren obere Hülle ebenfalls die Funktion $\underline{H}(P)$. Wir können uns daher im folgenden mit der Betrachtung der U.F. mit der unteren Schranke m begnügen. Aus 5. folgert man leicht, daß aus einer U.F. φ wieder eine solche, die mit φ_x bezeichnet werde, entsteht, wenn man φ in x durch $L_{\varphi, x}$ ersetzt. Wegen

$$m \leq \varphi \leq L_{\varphi, x}$$

ist also

$$(6, 2) \quad m \leq \varphi \leq \varphi_x.$$

Führen wir dies unter Benutzung eines festen Kreises κ für alle U.F. durch, so folgt, daß die obere Hülle der φ_x wiederum mit \underline{H} identisch ist. Aus (5) und (6) folgt

$$m \leq \varphi_x \leq M.$$

Wegen

$$\varphi_x(P) = L_{\varphi, x}(P)$$

folgt aus 2., daß $\underline{H}(P)$ in κ L -konvex ist. Da dies für jeden beliebigen Kreis aus \mathfrak{S} in Ω gilt, so haben wir das Resultat:

Die obere Hülle $\underline{H}(P)$ der U. F. ist L -konvex, mithin auch stetig.

Unser Ziel ist nun weiterhin, zu zeigen, daß sogar die Gleichung

$$(7, 2) \quad L_{\underline{H}, x}(P) = \underline{H}(P)$$

gilt, daß mit anderen Worten $\underline{H}(P)$ eine Lösung in Ω darstellt. Zum Beweis kann man ähnlich vorgehen wie in den Ausführungen von Brelot, nur ist jetzt zu berücksichtigen, daß in unserem Falle die Konstante nicht wie in der Potentialtheorie Lösung zu sein braucht.

Sei κ ein abgeschlossener Kreis aus \mathfrak{S} . Dann kann man leicht wie bei Brelot unter Benutzung des Borelschen Überdeckungssatzes eine U.F. φ bestimmen, welche in κ der Beziehung

$$(8, 2) \quad \varphi(P) \geq \underline{H}(P) - \varepsilon$$

genügt. Die U.F. $\varphi_x(P)$, die in κ gleich $L_{\varphi, x}(P)$ ist, befriedigt dann erst recht die Ungleichung (8). Da auch auf dem Rande von κ

$$\varphi = \varphi_x \geq \underline{H} - \varepsilon$$

ist, so folgt aus dem Maximumprinzip

$$L_{\varphi, x}(P) = \varphi_x(P) \geq L_{\underline{H}-\varepsilon, x}(P) = L_{\underline{H}, x}(P) - L_{\varepsilon, x}(P) \quad (P \in \kappa).$$

denn wegen der Stetigkeit von \underline{H} haben diese Operatoren einen Sinn. Mithin gilt in κ

$$\underline{H}(P) \geq \varphi_x(P) \geq L_{\underline{H}, x}(P) - L_{\varepsilon, x}(P).$$

Da wegen des Maximumprinzips

$$0 \leq L_{f,x}(P) \leq \varepsilon$$

ist, so folgt erst recht

$$\underline{H}(P) \geq L_{H,x}(P) - \varepsilon$$

für jedes ε . Also ist

$$(9, 2) \quad \underline{H}(P) \geq L_{H,x}(P).$$

Dies zusammen mit der L -Konvexität ergibt schließlich das gewünschte Resultat.

In ganz entsprechender Weise erhält man für die untere Hülle $\bar{H}(P)$ der O.F. ebenfalls die Gleichung

$$(7, 2) a \quad L_{\bar{H},x}(P) = \bar{H}(P).$$

Aus der Beziehung (2, 2) für die U.F. und der entsprechenden für die O.F. ergibt sich

$$\underline{\psi}(p) \geq f(p) \geq \overline{\varphi}(p),$$

$$\underline{\psi}(p) - \overline{\varphi}(p) \geq 0.$$

Für die untere Grenze $\underline{\psi}(P) - \overline{\varphi}(P)$ der Funktionen $\psi(P) - \varphi(P)$ im Randpunkt p folgt mithin

$$\underline{\psi}(p) - \overline{\varphi}(p) \geq \underline{\psi}(P) - \overline{\varphi}(P) \geq 0.$$

Da aber $\psi(P) - \varphi(P)$ L -konkav ist, so folgt aus 4. die Gültigkeit von

$$(10, 2) \quad \psi(P) - \varphi(P) \geq 0$$

in ganz Ω , und daraus

$$(11, 2) \quad \bar{H}(P) \geq \underline{H}(P)$$

in Ω .

Aus diesen Betrachtungen allein können wir zunächst noch nichts Näheres über \underline{H} und \bar{H} erschließen. Ja, wir müssen einräumen, daß unter Umständen \underline{H} und \bar{H} identisch Null sind.

Die Funktionen \underline{H} und \bar{H} sowie alle dazwischen liegenden Lösungen bezeichnen wir im folgenden einfach mit $u(P)$ oder $u(x, y)$. $\underline{H}(P)$ und $\bar{H}(P)$ nennen wir auch Hauptlösungen.

Für den Fall, daß Ω mit einem der Kreise κ aus \mathfrak{S} zusammenfällt, ist bei stetigem $f(p)$ $L_{f,x}(P)$ offenbar sowohl U.F. wie O.F. und mithin

$$\underline{H}(P) = \bar{H}(P) = L_{f,x}(P).$$

Wir bemerken noch, daß \underline{H} und \bar{H} und damit jede Lösung u überhaupt das folgende Maximumprinzip erfüllen

$$\sup |u(P)| \leq \sup |f(p)|.$$

denn es ist ja $\varphi(P) \equiv \sup |f(p)|$ eine O.F. und $\varphi(P) \equiv -\sup |f(p)|$ U.F. Ist jedoch bekannt, daß u nicht konstant ist, dann gilt sogar

$$|u(P)| < \sup |f(p)|,$$

wie man leicht aus IV folgert.

§ 3.

Verhalten auf dem Rande bei kleinen Gebieten.

Wir untersuchen nun, inwieweit die im vorigen Paragraphen konstruierten Funktionen $\underline{H}(P)$ und $\overline{H}(P)$ die folgende von Perron aufgestellte Randbedingung

$$(R) \quad l(p) \leq \underline{u}(p) \leq \overline{u}(p) \leq \overline{l}(p)$$

erfüllen; $\underline{u}(p)$ bzw. $\overline{u}(p)$ bedeutet den unteren bzw. oberen limes der in Ω definierten Funktion $u(P)$ im Randpunkte p .

In Anlehnung an entsprechende Begriffe in der Potentialtheorie führen wir folgende Definitionen ein.

Definition 4. p sei ein Randpunkt der in Ω_0 gelegenen Punktmenge Ω . Dann heißt p regulär in bezug auf Ω , wenn für jeden Durchschnitt $\Omega \kappa$ von Ω mit einem beliebigen Kreise κ um p und für jede beschränkte auf dem Rande von $\Omega \kappa$ definierte Funktion die zugehörigen Lösungen \underline{H} und \overline{H} in p die Randbedingung (R) erfüllen. Andernfalls heißt p singular.

Definition 5. Sei τ ein abgeschlossener Kreis um den Randpunkt p . Eine in $\Omega \tau$ definierte nichtnegative und stetige Funktion $w(P)$ heißt ein Barrier in p für τ bezüglich Ω , wenn sie in $\Omega \tau$ L -konkav ist, auf dem Rande von τ über einer positiven Schranke bleibt und in p den limes Null hat:

$$(1, 3) \quad \lim_{P \rightarrow p} w(P) = 0.$$

Wir beschränken unsere Betrachtung zunächst auf hinreichend kleine Mengen und beweisen:

Es sei p ein (innerer) Punkt des Definitionsgebietes Ω_0 der Operation L . Wenn für einen hinreichend kleinen Kreis τ_0 um p $L_{1, \tau_0}(p) > 0$ ist, so ist p für die in τ_0 gelegene offene Punktmenge Ω , zu deren Rande p gehört, regulär, wenn für jeden in τ_0 liegenden Kreis τ um p ein Barrier in bezug auf Ω existiert.

Wir bemerken zunächst, daß man jedes Barrier über ganz Ω L -konkav erweitern kann. Denn es sei etwa auf dem Rande von τ

$$(2, 3) \quad w(P) \geq 1 + \eta \quad (\eta > 0).$$

Dann ist die Funktion

$$(3, 3) \quad \begin{aligned} w^*(P) &= \text{Min} (1, w(P)) && \text{in } \Omega \tau \\ &= 1 && \text{in } \Omega - \Omega \tau \end{aligned}$$

L -konkav in Ω .

Beweis. Die Behauptung gilt innerhalb $\Omega \tau$ wegen 3., innerhalb $\Omega - \Omega \tau$ wegen des Maximumprinzips. Ist nun P ein beliebiger Punkt von Ω , so kann ich, falls P auf dem Rande von τ liegt, wegen der Stetigkeit von $w(P)$ auf dem Rande von τ , und wegen (2, 3) einen so kleinen Kreis κ um P bestimmen, daß in ihm noch $w(P) > 1$ ist, so daß in κ

$$w^*(P) = 1$$

ist; dann gilt auch hier noch

$$w^*(P) \geq L_{w^*, \kappa}(P),$$

w. z. b. w. Ist nun

$$\overline{f(p)} \geq 0,$$

so kann man fast genau wie bei Perron zeigen, daß die Randbedingung

$$\overline{u(p)} \leq \overline{f(p)}$$

erfüllt ist. Da nämlich jede positive Konstante L -konkav ist, so ist auch die Funktion

$$\overline{f(p)} + \varepsilon + C \cdot w^*(P) = \varphi(P) \quad (C > 0)$$

L -konkav, und, wenn die Konstante C hinreichend groß gewählt ist, eine O.F. Nun ist wegen der Eigenschaften von w^*

$$\lim_{P \rightarrow p} \varphi(P) = \overline{f(p)} + \varepsilon$$

und es gibt eine Umgebung U von p derart, daß in U

$$\varphi(P) \leq \overline{f(p)} + 2\varepsilon,$$

also erst recht

$$u(P) \leq \overline{f(p)} + 2\varepsilon$$

ist. Mithin gilt

$$\overline{u(p)} \leq \overline{f(p)} + 2\varepsilon$$

für jedes ε und es ergibt sich schließlich

$$(4, 3) \quad \overline{u(p)} \leq \overline{f(p)}.$$

In gleicher Weise folgert man aus

$$\underline{f(p)} \leq 0$$

die Ungleichung

$$(4, 3) a \quad \underline{u(p)} \geq \underline{f(p)}.$$

Wenn also gleichzeitig

$$\underline{f}(p) \leq 0 \leq \overline{f}(p)$$

ist, so ist hiermit das Erfülltsein der Randbedingung bewiesen.

Sei jetzt etwa

$$0 < \underline{f}(p) \leq \overline{f}(p).$$

Dann ist, wie soeben bewiesen,

$$\overline{u}(p) \leq \overline{f}(p).$$

Um die andere Hälfte der Randbedingung zu beweisen, betrachten wir das Problem mit der Randfunktion

$$g(q) = f(q) - C \cdot L_{1,\tau_0}(q),$$

die ja einen Sinn hat, weil Ω in τ_0 liegt. Nach Voraussetzung ist

$$(5, 3) \quad L_{1,\tau_0}(p) = \eta > 0.$$

Die Konstante C wählen wir so groß, daß

$$C \cdot \eta > \underline{f}(p)$$

und mithin

$$(6, 3) \quad g(p) = \underline{f}(p) - C\eta < 0$$

ist. Wegen (6, 3) ist für die zugehörigen Lösungen $u_g(P)$ die Randbedingung

$$(7, 3) \quad \underline{u}_g(p) \geq g(p) = \underline{f}(p) - C\eta$$

erfüllt. Nun ist aber für jede zu f gehörige Lösung u

$$u(P) - C \cdot L_{1,\tau_0}(P)$$

eine Lösung u_g . Denn da $L_{1,\tau_0}(P)$ in Ω einschließlich des Randes stetig ist, so ist, wenn z. B. $\psi(P)$ eine O.F. von $u_g(P)$ ist, offenbar

$$\psi(P) + C \cdot L_{1,\tau_0}(P)$$

eine O.F. von u . Entsprechendes gilt für Unterfunktionen.

Aus der Gleichung

$$u(P) = u_g(P) + C \cdot L_{1,\tau_0}(P)$$

folgt aber sofort mittels (7, 3)

$$(8, 3) \quad \underline{u}(p) = \underline{u}_g(p) + C \cdot L_{1,\tau_0}(p) \geq \underline{f}(p),$$

w. z. b. w.

Wir wollen jetzt Entsprechendes auch im großen durchführen.

§ 4.

Verhalten auf dem Rande bei beliebigen Gebieten.

Satz 1. p sei ein Punkt von Ω_0 und Randpunkt der in Ω_0 gelegenen offenen Punktmenge Ω . Für einen hinreichend kleinen Kreis τ_0 um p sei

$$L_{1,\tau_0}(p) > 0.$$

Dann ist der Randpunkt p regulär in bezug auf Ω , wenn für jeden hinreichend kleinen Kreis ein Barrier bezüglich p existiert.

Beweis. S sei der Rand von Ω , s_0 der Rand von τ_0 . Nach dem vorher Bewiesenen ist p regulär in bezug auf die Punktmenge $\Omega\tau_0$. Jetzt sei $\psi(P)$ eine O. F. für Ω und die Randfunktion $f(q)$. Wir schreiben nun auf dem Rande von $\Omega\tau_0$ folgende Randwerte $g(r)$ vor

$$g(r) = \psi(r) + \varepsilon \quad (\varepsilon > 0, r \text{ ein Punkt des Durchschnitts } s_0\Omega),$$

$$g(q) = f(q) \quad (q \text{ ein Punkt des Durchschnitts } (\tau_0 + s_0)S),$$

speziell also

$$g(p) = f(p),$$

$$\underline{g(p)} = \underline{f(p)}, \quad \overline{g(p)} = \overline{f(p)}.$$

Wir konstruieren in Ω eine neue O. F. $\psi_{\tau_0}(P)$ in folgender Weise:

$\psi_{\tau_0}^*(P)$ sei eine O. F. für $\Omega\tau_0$. Dann setzen wir

$$\begin{aligned} \psi_{\tau_0}(P) &= \text{Min}(\psi_{\tau_0}^*, \psi) & (P \in \Omega\tau_0) \\ &= \psi(P) & (P \in \Omega - \Omega\tau_0). \end{aligned}$$

Auf $s_0\Omega$ gilt

$$\underline{\psi_{\tau_0}^*}(r) \geq \psi(r) + \varepsilon.$$

In einer gewissen Umgebung von r gilt also noch

$$\psi_{\tau_0}^*(P) > \psi(r).$$

Dies gilt also auch in einer gewissen Umgebung U_0 von ganz $s_0\Omega$, soweit sie zu $\Omega\tau_0$ gehört. In dieser ist folglich noch

$$\psi_{\tau_0}(P) = \psi(P).$$

Aus der Stetigkeit von ψ und $\psi_{\tau_0}^*$ in $\Omega\tau_0$ folgt somit auch die Stetigkeit von $\psi_{\tau_0}(P)$ in Ω . Wir haben noch die Konkavität von ψ_{τ_0} zu beweisen. Dies gelingt ähnlich wie vorhin die Erweiterung des Barriers über ganz Ω (S. 741). Innerhalb $\Omega\tau_0$ und $\Omega - \Omega\tau_0$ folgt die Konkavität aus 3. bzw. aus der Tatsache, daß in $\Omega - \Omega\tau_0$ $\psi_{\tau_0} = \psi$, und ψ O. F. ist. Die Kreise mit den Mittelpunkten auf s_0 wähle man so klein, daß sie noch in der Umgebung U_0 liegen; in dieser ist aber, wie vorhin bemerkt,

$$\psi_{\tau_0} = \psi$$

und die Konkavität also ebenfalls gesichert. ψ_{τ_0} ist also, wie behauptet, in Ω L -konkav.

Da nun $\psi_{\tau_0}^*$ O.F. in Ω_{τ_0} , und ψ O.F. in Ω ist, so gilt sowohl auf $S(\tau_0 + s_0)$ wie auf $S - S(\tau_0 + s_0)$ die Ungleichung

$$\psi_{\tau_0}(q) \geq f(q).$$

$\psi_{\tau_0}(P)$ ist also in der Tat O.F. für Ω und die Randwerte $f(q)$. Wegen der Regularität von p für Ω_{τ_0} kann man die O.F. $\psi_{\tau_0}^*$ so wählen, daß

$$\overline{\psi_{\tau_0}^*}(p) \leq \overline{f}(p) + \varepsilon,$$

also auch

$$\overline{\psi_{\tau_0}}(p) \leq \overline{f}(p) + \varepsilon$$

ist. Sei $\overline{H}(P)$ die untere Hülle der O.F. in Ω ; dann ergibt sich schließlich

$$\overline{H}(p) \leq \overline{f}(p).$$

In völlig entsprechender Weise läßt sich zeigen

$$\underline{H}(p) \geq \underline{f}(p).$$

Damit ist der Satz vollständig bewiesen.

Gleichzeitig ist damit noch gezeigt:

Aus der Regularität im kleinen folgt die Regularität im großen.

§ 5.

Bedingung VI. Klassische Lösung der Randwertaufgabe.

Wir fügen den Bedingungen I—V eine weitere hinzu, die von jetzt ab durchweg in Geltung bleiben soll.

VI^a). Konvergenzbedingung. Ist κ ein beliebiger Kreis aus \mathfrak{S} und konvergieren die gleichmäßig beschränkten Funktionen $f_n(p)$ außer in endlich vielen Punkten gegen Null, so konvergieren in jedem inneren Punkte von κ auch die Funktionen $L_{n,\kappa}(P)$ gegen Null. Außerdem sei für jeden Kreis aus \mathfrak{S}

$$L_{1,\kappa}(P) > 0.$$

Aus VI folgt speziell, daß man den Wert einer Funktion $L_{f,\kappa}(P)$, deren Betrag 1 nicht überschreitet, für jeden inneren Punkt beliebig klein machen kann, wenn man ihre Randwerte auf einem hinreichend langen Randbogen gleich Null macht. Zieht man eine solche Funktion von $L_{1,\kappa}(P)$ ab, so kann man erreichen, daß die Funktion $L_{f,\kappa}(P)$ im Mittelpunkt des Kreises größer als

^a) VI ließe sich noch in verschiedener Weise abschwächen. Aus Gründen der Übersichtlichkeit ist jedoch die obige Formulierung gewählt worden.

Null ist, wenn sie nur absolut unter einer festen Schranke bleibt, und wenn ihre Randwerte f auf einem hinreichend langen Peripheriebogen größer oder gleich 1 sind.

Weiterhin ergibt sich leicht die eindeutige Lösbarkeit des Randwertproblems für stückweise stetige Funktionen f auf dem Rande der Kreise κ . Man kann eine solche leicht von oben oder von unten her durch stetige Funktionen approximieren. Wegen VI strebt die Differenz der zugehörigen O.F. und U.F. gegen Null; und man erkennt, daß die Lösung H außer an den Sprungstellen die Randwerte f besitzt. Dagegen weiß ich nicht, ob man umgekehrt aus der eindeutigen Lösbarkeit für stückweise stetige Funktionen auch die Eigenschaft VI folgern kann.

Durch die Erweiterung der Operation $L_{f,\kappa}$ für stückweise stetige Funktionen ist nun ein lineares Funktional über dem Raume C aller stetigen Funktionen $f(p)$ der Strecke $0 \leq s \leq l$ (l Länge des Kreisumfanges von κ) definiert mit der Norm $\|f\| = \text{Max } |f(s)|$. Also ist $L_{f,\kappa}$ mittels einer Funktion von beschränkter Schwankung $v(s)$ in der Form

$$L_{f,\kappa}(P) = \int_0^l f(p) dv_P(p)$$

darstellbar. Wir wollen nun folgendes zeigen:

Ist $\{f(s)\}$ eine Schar von gleichmäßig beschränkten stetigen Funktionen, die in einem festen Punkte der Peripherie gleichgradig stetig sind, so sind auch die zugehörigen Funktionen $L_{f,\kappa}(P)$ in demselben Randpunkte gleichgradig stetig.

Beweis. Wir zerlegen die Peripherie l in zwei zusammenhängende Stücke s' und s'' . Für einen inneren Punkt p_0 von s' ist offenbar $\lim_{P \rightarrow p_0} v_P(s') = 1$, $\lim_{P \rightarrow p_0} v_P(s'') = 0$. Es gibt also einen Kreis von hinreichend kleinem Radius δ um p_0 derart, daß $v_P(s'') < \varepsilon$ bleibt, wenn $\overline{p_0 P} < \delta$ ist. Nun seien die Funktionen $f(p)$ stetig und gleichmäßig beschränkt: $|f(p)| < M$. Wir können uns auf den Fall beschränken, daß für alle Funktionen $f(P)$

$$f(p_0) = 0$$

ist. Diese seien außerdem gleichgradig stetig in p_0 . Es gibt also eine Umgebung s' von p_0 auf l derart, daß für alle f

$$|f(p) - f(p_0)| = |f(p)| < \varepsilon$$

gilt, falls p auf s' liegt. Dann folgt aus $\overline{p_0 P} < \delta$

$$\left| \int_{s'} f(p) dv_P(p) \right| < \varepsilon v_P(s') \leq \varepsilon, \quad \left| \int_{s''} f(p) dv_P(p) \right| < M v_P(s'') < M \varepsilon,$$

$$\left| \int_l f(p) dv_P(p) \right| \leq \left| \int_{s'} \right| + \left| \int_{s''} \right| < \varepsilon (1 + M).$$

w. z. b. w.

Wir beweisen jetzt die folgenden beiden Sätze:

Satz 2. *Ist der Randpunkt p der offenen Punktmenge $\Omega \subset \Omega_0$ auch Randpunkt eines Kreises σ , der außerhalb Ω liegt, so ist p regulär in bezug auf Ω .*

Satz 3. *Die Barrierbedingung von Satz 1 ist nicht nur hinreichend, sondern auch notwendig für die Regularität eines Randpunktes.*

Aus Satz 2 folgt speziell die klassische Lösbarkeit des Randwertproblems für Gebiete Ω , die von endlich vielen stetig gekrümmten Kurven berandet werden, wobei noch Ecken und Spitzen zugelassen sind, falls die überstumpfe Seite des Winkels außerhalb Ω liegt. Wegen IV kann es natürlich nur eine Lösung geben.

Beweis von Satz 1. Da wegen VI jetzt immer $L_{1,z}(P) > 0$ ist, so haben wir nur die Existenz eines Barriers nachzuweisen. Die Konstruktion eines solchen ist allerdings reichlich umständlich. Den gemäß Definition 5 heranzuziehenden Kreis τ_0 mit dem Mittelpunkt p wählen wir so klein, daß er den im Satz erwähnten Kreis σ in zwei Punkten schneidet. Die Verlängerung der Zentrale der beiden Kreise möge den außerhalb σ liegenden Bogen des Randes von τ_0 in Q schneiden. Auf dem Rande von τ_0 wird nun in folgender Weise eine Funktion $g(s)$ definiert: Q'_1, Q'_1 und Q'_2, Q'_2 seien zwei verschiedene Punktepaare, die symmetrisch zu Q auf dem Rande s_{τ_0} von τ_0 liegen; Q'_1, Q'_1 liegen näher an Q . s_0 sei der von ihnen begrenzte Teilbogen von s_{τ_0} , welcher Q enthält. s_1 sei der von dem anderen Paar begrenzte Teilbogen, der Q nicht enthält. Dann sei

$$\begin{aligned} g(s) &= \varepsilon > 0 \quad \text{auf } s_0 \\ &= -1 \quad \text{auf } s_1. \end{aligned}$$

In den zwischen Q'_1, Q'_2 bzw. Q'_1, Q'_2 liegenden Restbögen ergänzen wir $g(s)$ linear. Die Länge von s_0 und s_1 und die Größe von ε sei nach der Bemerkung zur Bedingung VI so gewählt, daß

$$L_{\sigma, \tau_0}(p) < 0$$

ist. Dann gibt es einen Kreis τ um p , in welchem ebenfalls noch

$$(1, 5) \quad L_{\sigma, \tau_0}(P) < 0 \quad (P \in \tau)$$

ist. Nun sei τ' ein Kreis durch p , welcher den Kreis τ_0 in einem Punkte $Q' \neq Q$ des Bogens $\overline{QQ'_1}$ berühre. Q_0 sei sein Schnittpunkt mit der Geraden \overline{Qp} . Wir wählen Q' so nahe bei Q , daß auf dem Bogen $\widehat{Q'Q_0}$ von $s_{\tau'}$ noch

$$(2, 5) \quad L_{\sigma, \tau_0}(P) > 0$$

bleibt. τ'' sei der zu τ' symmetrisch liegende Kreis. Setzen wir

$$(3, 5) \quad Cr_{pq} = h(q) \quad (r_{pq} = \overline{pq}),$$

wobei die Konstante C so groß gewählt sei, daß, wenn ϱ den Radius des kleinen Kreises τ um p bedeutet,

$$C\varrho > \varepsilon > 0$$

ist. Da ε das Maximum von $L_{\varrho, \tau_0}(P)$ ist, so ist mithin außerhalb τ

$$(4, 5) \quad h(q) > \varepsilon \geq L_{\varrho, \tau_0}(P).$$

Nun bestimmen wir in $\tau' + \tau''$ in folgender Weise eine L -konkave Funktion $u(P)$:

Zunächst definieren wir durch einen alternierenden Prozeß in jedem der Kreise τ' und τ'' eine Folge von Funktionen $\{u'_i\}$, $\{u''_i\}$:

$$\begin{aligned} u'_1 &= L_{h'_1, \tau'} & h'_1 &= h \text{ auf } s_{\tau'}, \\ u''_1 &= L_{h''_1, \tau''} & h''_1 &= h \text{ außerhalb } \tau' \\ & & &= u'_1 \text{ innerhalb } \tau', \\ u'_2 &= L_{h'_2, \tau'} & h'_2 &= h \text{ außerhalb } \tau'' \\ & & &= u''_1 \text{ innerhalb } \tau'', \\ & \dots \dots \dots \end{aligned}$$

Wegen (4, 5) ist

$$(5, 5) \quad \begin{aligned} u'_i(P) &\geq L_{\varrho, \tau_0}(P), \\ u''_i(P) &\geq L_{\varrho, \tau_0}(P), \end{aligned} \quad (i = 1, 2, \dots).$$

Außerdem gilt auch noch

$$(6, 5) \quad \begin{aligned} u'_i(P) &\geq u'_0(P), \\ u''_i(P) &\geq u''_0(P), \end{aligned}$$

Dabei ist z. B. $u'_0(P) = L_{\varrho_0, \tau'}(P)$, wenn g'_0 auf den in τ'' liegenden Randpunkten von τ' gleich 0 ist, und auf dem Restbogen von $s_{\tau'}$ gleich $h(q)$, mit Ausnahme eines beliebig kleinen Bogens bei Q_0 , auf welchem g'_0 durch eine lineare Funktion stetig ergänzt wird. $u''_0(P)$ ist symmetrisch definiert.

Nun sei $u'(P)$ die untere Hülle der u'_i , $u''(P)$ die untere Hülle der u''_i .

Wir setzen

$$\begin{aligned} u(P) &= u'(P) & (P \in \tau' - \tau' \tau'') \\ &= u''(P) & (P \in \tau'' - \tau' \tau'') \\ &= \text{Min}(u', u'') & (P \in \tau' \tau''). \end{aligned}$$

Die Kreise werden dabei als offene Punktmengen betrachtet. Da, wie man leicht sieht, auf dem Rande von $\tau' \tau''$ $\text{Min}(u', u'')$ gleich u' in den inneren Punkten von τ' und gleich u'' in den inneren Punkten von τ'' ist, so ist $u(P)$ also in jedem Punkt die untere Grenze der in ihm definierten Funktionen u'_i , u''_i . Wir behaupten:

u ist stetig in der abgeschlossenen Hülle von $\tau' + \tau''$ mit möglicher Ausnahme des Punktes Q_0 .

Wegen V sind nämlich die u'_i in jedem inneren Punkte von τ' gleichgradig stetig. Daher sind die Randwerte aller u'_i in den inneren Punkten des in τ' liegenden Randbogens von τ'' gleichgradig stetig, und daher auch nach den Bemerkungen zur Bedingung VI die u'_i selbst. Daraus folgt die Stetigkeit von u im Innern von $\tau' + \tau''$. Nun ergibt sich aus (6, 5)

$$u'_1 \geq u \geq u'_0,$$

$$u''_1 \geq u \geq u''_0.$$

Man kann u'_0 so bestimmen, daß in einem beliebigen inneren Punkte q des äußeren Bogens $p\bar{Q}_0$ von $s_{\tau'}$, u'_0 den Randwert $h(q)$ annimmt; auch $u'_1(q)$ ist gleich $h(q)$, also nimmt u in q stetig den Wert $h(q)$ an. Das gleiche gilt selbstverständlich auch in p . Damit ist die Behauptung bewiesen.

Wir behaupten weiter

u ist L -konkav in $\tau' + \tau''$.

Dies bedarf wegen 3. nur für die inneren Randpunkte der Kreise τ' und τ'' eines Beweises. κ sei ein Kreis um einen in τ' gelegenen Randpunkt von τ'' . Dann ist

$$L_{u,\kappa}(P) \leq L_{u',\kappa}(P) = u'_i(P),$$

mithin

$$L_{u,\kappa}(P) \leq \inf u'_i(P) = u'(P).$$

Also ist

$$L_{u,\kappa}(P) \leq u(P) \quad (P \in \tau' - \tau'').$$

Bleibt noch zu zeigen, daß in $\tau' \tau''$ auch

$$L_{u,\kappa}(P) \leq u''(P)$$

ist. Bedeuten $s_{\tau''}$ und s_{κ} den Rand von τ'' bzw. κ , so ist auf $s_{\tau''}$ $u'' = u' = u$, mithin $L_{u,\kappa}(P) \leq u = u''$. Auf $s_{\kappa} \tau''$ ist $L_{u,\kappa}(P) = u \leq u'$; also ist auf dem ganzen Rande des Durchschnitts $\kappa \tau'' = \omega$: $L_{u,\kappa} \leq u''$. Setzen wir $L_{u,\kappa} = v$, $u'' - v = w$, so gilt, da wegen 3. für alle Kreise τ in τ'' $L_{u'',\tau} \leq u''$ ist, $L_{w,\tau} \leq w$ in ω ($\tau \subset \omega$), während auf dem Rande $w \geq 0$ ist. Nach 4. ist die in ω L -konkave Funktion w auch im Innern von ω nichtnegativ, d. h. es ist, wie behauptet,

$$u'' \geq v = L_{u,\kappa} \quad (P \in \omega).$$

Aus (5, 5) folgt nun

$$(7, 5) \quad u(P) \geq L_{g,\tau_0}(P)$$

und aus (2, 5)

$$L_{g,\tau_0}(Q_0) > 0.$$

Wegen der Stetigkeit gibt es also einen hinreichend kleinen Kreis τ_1 um Q_0 und eine positive Konstante η derart, daß

$$L_{g,\tau_0}(P) > \eta > 0$$

gilt für alle Punkte des Kreises τ_1 . Aus (7, 5) folgt schließlich

$$(8, 5) \quad u(P) > \eta \quad (P \in \tau_1(\tau' + \tau''))$$

auf der Kontur γ des Gebietes $\tau' + \tau'' - \tau_1(\tau' + \tau'')$, auf welcher u noch stetig ist, und außerhalb τ gilt also wegen (4, 5) und (8, 5)

$$(9, 5) \quad u(P) \geq \text{Min}(\varepsilon, \eta) = \eta_1 > 0.$$

Nun schneidet der Kreis σ wegen der Lage von τ' diesen Kreis außer in dem Punkte p sicher noch in einem weiteren Punkte p' . Wir können den Kreis τ so klein gewählt denken, daß er den Punkt p' nicht enthält. Dann gilt die Beziehung (9, 5) also auf dem außerhalb σ liegenden Teile von γ , also erst recht in den in $\Omega\tau_0$ liegenden Stücken von γ . Setzen wir nun

$$\begin{aligned} w(P) &= \text{Min} \left(u(P), \frac{\eta_1}{2} \right) \text{ in } \Omega \cdot (\tau' + \tau'' - \tau_1(\tau' + \tau'')) \\ &= \frac{\eta_1}{2} \quad \text{im Rest von } \Omega\tau_0, \end{aligned}$$

so ist $w(P)$ in $\Omega\tau_0$ L -konkav, wie man ganz ähnlich wie in früheren Fällen beweist.

$w(P)$ ist mithin ein Barrier bezüglich Ω in p und p folglich regulär, wie im Satz behauptet wurde.

Satz 2 ermöglicht sofort den Beweis des Satzes 3. Ist nämlich p regulär, so ist in $\Omega\tau$ (τ wieder ein Kreis um p) jede zu den Randwerten $f(q) = \overline{p}q$ gehörige Lösung ein Barrier. Denn die in Ω liegenden Randpunkte von τ sind regulär, so daß die Lösungen dort die Randwerte $\overline{p}q = \text{konst} > 0$ annehmen.

Wir können jetzt die Definition 4 in § 3 dahingehend abändern, daß wir im Wortlaut statt der beschränkten Randfunktionen f nur stetige benutzen. Denn es genügt ja für die Existenz eines Barrier, wenn für die eben benutzte Funktion $f(q) = \overline{p}q$ die Randbedingung erfüllbar ist. Außerdem genügt es, die Erfüllung der Randbedingung (R) nur für \underline{H} oder \overline{H} allein zu fordern, weil man trotzdem immer ein Barrier konstruieren kann. Es gilt also z. B.:

Erfüllt in allen $\Omega\tau$ für alle stetigen Randfunktionen \underline{H} die Randbedingung in p , so wird sie auch von \overline{H} befriedigt und sie ist dann sogar für alle beschränkten Randfunktionen erfüllt.

2. Das verallgemeinerte Problem.

§ 6.

Normierte Potentiale.

ω_0 sei ein Gebiet, dessen Rand s_0 nur aus regulären Punkten besteht, so daß die Randwertaufgabe klassisch lösbar ist. $\bar{\omega}$ mit dem Rande \bar{s} sei eine abgeschlossene Teilmenge von ω_0 . Die Differenzmenge $\omega_0 - \bar{\omega}$ zerfällt in Gebiete, von denen eines, das mit ω^* bezeichnet werde, den Rand s_0 von ω_0

als Randkomponente besitzt. Der Rest des Randes von ω^* sei s_1 . Für das Gebiet ω^* lösen wir nun die Randwertaufgabe mit folgender Randvorgabe

$$\begin{aligned} h(p) &= 0 & (p \in s_0) \\ &= 1 & (p \in s_1). \end{aligned}$$

Wie sich nachher zeigen wird, sind die zugehörigen Hauptlösungen \underline{H} und \bar{H} identisch. Die Funktion

$$u(P) = \underline{H}(P) = \bar{H}(P)$$

bezeichnen wir als das *normierte Potential*⁷⁾ von $\bar{\omega}$ für das Gebiet ω_0 . Die Bezeichnung Potential wird erst im nächsten Paragraphen ihre Begründung finden.

Um zu zeigen, daß $\underline{H} = \bar{H}$ ist, approximieren wir $\omega_0 - \omega^*$ durch ineinandergeschachtelte Mengen $\omega_n \supset \omega_{n+1} \supset \omega_0 - \omega^*$, die von s_0 und analytischen Kurven s_n ⁸⁾ begrenzt werden. Die zugehörigen normierten Potentiale seien $u_n(P)$. Diese erfüllen überall die Randbedingung, nehmen also auf s_n die Werte 1 an. Die Funktionen

$$\begin{aligned} \psi_n(P) &= u_n(P) & (P \in \omega_0 - \omega_n) \\ &= 1 & (P \in \omega_n) \end{aligned}$$

können daher als O.F. für \bar{H} benutzt werden. Die L -Konkavität läßt sich ohne Schwierigkeit beweisen. Die Funktionen sind außerdem monoton abnehmend und konvergieren wegen der gleichgradigen Stetigkeit in $\omega^* + s_0$ gegen eine stetige Funktion u^* . Wegen der Monotonie ist die Konvergenz in jedem abgeschlossenen Kreise τ in ω^* sogar gleichmäßig. Hieraus folgert man leicht auf Grund des Maximumprinzips, daß in τ

$$u^{(n)}(P) = L_{u^{(n)}, \tau}(P)$$

gegen $L_{u^*, \tau}(P)$ konvergiert. Es ist also

$$u^*(P) = L_{u^*, \tau}(P),$$

für jeden Kreis τ aus $\bar{\omega}$ in ω^* . Auf s_1 gilt

$$\overline{u^*(p)} \leq h(p).$$

$u^*(P)$ ist also U.F., d. h. aber \underline{H} und \bar{H} stimmen überein und fallen mit u zusammen wie behauptet:

$$u(P) = u^*(P) = \underline{H}(P) = \bar{H}(P).$$

⁷⁾ Bei dieser Definition, die für unseren Zweck vollständig ausreicht, kommt nur der Außenrand s_1 der Menge $\bar{\omega}$ ins Spiel. Die Menge selbst ist also noch weitgehend unbestimmt.

⁸⁾ Vgl. O. D. Kellogg, Foundations of Potential Theory, S. 319.

Es sei jetzt Ω irgendeine offene Punktmenge in Ω_0 , p ein Randpunkt von Ω und τ_1 ein abgeschlossener Kreis um p , $\bar{\omega}$ der abgeschlossene Durchschnitt von τ_1 mit der Komplementärmenge von Ω . τ_0 sei ein größerer konzentrischer Kreis um p

$$\tau_0 \supset \tau_1.$$

Satz 4. *Dann ist p regulär oder singulär, je nachdem ob für alle hinreichend kleinen Kreise τ_1 das normierte Potential $u(P)$ von $\bar{\omega}$ für τ_0 in p den Randwert 1 annimmt oder nicht.*

Denn offenbar ist die Funktion $1 - u(P)$ ein Barrier bezüglich Ω τ_0 .

Damit haben wir ein zweites und für die Folge wichtiges Kriterium für die Regularität gewonnen.

§ 7.

Greensche Funktion. Kapazität.

Wir müssen jetzt noch zwei Bedingungen einführen, die von nun an dauernd in Geltung bleiben sollen.

VII. Existenz einer Grundlösung. In jedem Kreise κ aus \mathfrak{S} existiere zu jedem inneren Punkte Q von κ eine Funktion $F(P, Q)$, die folgende Eigenschaften besitzt:

1) Bei festem Q ist $F(P, Q)$, falls $P \neq Q$, eine Lösung, d. h. wenn $\tau \subset \kappa$ den Punkt Q , den Pol, weder in seinem Innern noch auf dem Rande enthält, ist

$$F(P, Q) = L_{F, \tau}(P).$$

2) $F(P, Q) = \lg \frac{1}{r_{PQ}} + w(P, Q)$. Bei festem $P \neq Q$ ist $w(P, Q)$ stetig in Q). Wird Q auf eine abgeschlossene Punktmenge α in κ beschränkt, die also vom Rande von κ einen positiven Abstand besitzt, so gibt es eine endliche Zahl M derart, daß

$$|w(P, Q)| < M \quad (P \in \kappa, Q \in \alpha)$$

ist.

Addiert man zur Grundlösung die Funktion

$$-L_{F, \kappa}(P) = - \int F(R, Q) d\nu_P(s_R),$$

so erhält man eine auf dem Rande von κ verschwindende Grundlösung, d. h. eine *Greensche Funktion*

$$G(P, Q) = F(P, Q) - \int F(R, Q) d\nu_P(s_R),$$

welche wegen IV nichtnegativ sein muß.

²⁾ Daß auch das Umgekehrte gilt, folgt aus 1).

Sei jetzt σ eine abgeschlossene Punktmenge im Innern von κ , und $\mu(\omega)$ eine absolut additive Mengenfunktion auf σ . Dann existiert wegen der Stetigkeit von $G(P, Q)$ das Riemann-Stieltjes-Integral

$$(1, 7) \quad U(P) = \int_{\sigma} G(P, Q) d\mu(Q),$$

falls P nicht auf σ liegt, und stellt daselbst offenbar eine stetige Funktion dar. Wir wollen zeigen, daß $U(P)$ sogar eine Lösung ist.

Es ist zu zeigen

$$U(P) = L_{U, \tau}(P) \quad \tau \subset \kappa - \sigma.$$

Setzen wir

$$!G(P, Q_i) = G_i, \quad \mu^*(\omega_i) = \mu_i, \quad (Q_i \in \omega_i, \mu^*(\omega) = \mu(\omega\sigma))$$

und

$$\sum_i G_i \mu_i = U_n(P).$$

Dann ist

$$(2, 7) \quad U_n(P) = L_{U_n, \tau}(P).$$

Da τ von σ einen positiven Abstand hat, folgt aus Bedingung VII, 2) leicht die gleichmäßige Beschränktheit von $U_n(P)$ auf τ und daraus wegen V die gleichgradige Stetigkeit, und zwar im abgeschlossenen Kreise τ . Die $U_n(P)$ sind also auf τ sogar gleichmäßig stetig. Da sie in jedem Punkte von τ konvergieren, sobald die Durchmesser der ω_i gleichmäßig gegen Null streben, so konvergiert die Folge $\{U_n(P)\}$ sogar gleichmäßig auf τ . Daraus wiederum ergibt sich unter Benutzung des Maximumprinzips

$$L_{U_n, \tau}(P) \rightarrow L_{U, \tau}(P)$$

und wegen (2, 7) schließlich

$$U(P) = L_{U, \tau}(P).$$

Nach diesen Vorbereitungen stellen wir nun noch die folgende und letzte Forderung an die Operation L :

VIII. Darstellung des normierten Potentials. Die im Innern eines beliebigen Kreises κ aus \mathfrak{S} gelegene abgeschlossene Punktmenge $\bar{\omega}$ sei von endlich vielen analytischen Kurven berandet. Dann läßt sich das normierte Potential $u(P)$ von $\bar{\omega}$ mittels einer positiv monotonen Funktion $\mu(\omega)$ in der Form (1, 7) darstellen, wobei $\sigma = \bar{\omega}$ zu setzen ist.

Hiermit findet die Bezeichnung „Potential“ eine gewisse Rechtfertigung.

Es ist möglich, daß mehrere Normierungsbelegungen $\mu(\omega)$ existieren. Unter diesen greifen wir irgendeine heraus und bezeichnen die totale Variation $\mu(\bar{\omega}) = \gamma$ als die *Kapazität* von $\bar{\omega}$.

§ 8.

Vergleichung mit der Kapazität bei Δu .

Wir kommen zum Kernpunkt der Theorie. Es handelt sich darum, die oben definierte Kapazität für den Operator L mit der Kapazität bei Δu zu vergleichen. Dies wird uns in Verbindung mit dem nachher herangezogenen Wienschen Kriterium in den Stand setzen, die Gleichheit der regulären Punkte beim Operator L mit denen bei Δu zu beweisen.

Unter Beibehaltung der Bezeichnung von § 6 sei jetzt $\omega_0 = \kappa$ ein Kreis aus \mathfrak{S} . Die abgeschlossene Punktmenge $\bar{\omega}$ werde von analytischen Kurven berandet. $u(P) = \int_{\bar{\omega}} G(P, Q) d\mu(Q)$ sei das normierte Potential von $\bar{\omega}$. Die Belegungsfunktion $\mu(\omega)$ sei so gewählt, daß

$$(1, 8) \quad \mu(\bar{\omega}) = \gamma$$

ist. Dann existiert bekanntlich eine Gleichgewichtsverteilung $\mu^*(\omega)$ auf $\bar{\omega}$, für welche das Potential

$$(2, 8) \quad u^*(P) = \int_{\bar{\omega}} \lg \frac{1}{r_{PQ}} d\mu^*(Q)$$

auf $\bar{\omega}$ den Wert 1 besitzt. Wir schreiben

$$(3, 8) \quad \mu^*(\bar{\omega}) = \gamma^* \quad (\text{Kapazität für } \Delta u).$$

$\mu^*(\omega)$ ist bekanntlich in diesem Falle nur eine Linienbelegung, und zwar auf dem Außenrande von $\bar{\omega}$. Sie ist positiv, wenn, wie wir annehmen wollen, $\bar{\omega}$ einen hinreichend kleinen Durchmesser besitzt. Setzen wir

$$(4, 8) \quad \frac{\gamma}{\gamma^*} = \lambda.$$

Da die abgeschlossene Punktmenge $\bar{\omega}$ im Innern von κ liegt, so liegt sie auch im Innern eines zu κ konzentrischen kleineren Kreises τ . Gemäß VII, 2) gibt es also eine nur von τ abhängige Zahl M derart, daß, wenn

$$G(P, Q) = \lg \frac{1}{r_{PQ}} + w(P, Q)$$

gesetzt wird,

$$|w(P, Q)| < M$$

ausfällt, wie auch P in κ und Q in τ gewählt sein möge. Nehmen wir nun τ so klein an, daß

$$\frac{1}{2} \lg \frac{1}{r_{PQ}} \geq M > 0 \quad (P \in \tau, Q \in \tau)$$

gilt, so ist mithin

$$(5, \quad \frac{1}{2} \lg \frac{1}{r} < G(P, Q) < \frac{3}{2} \lg \frac{1}{r},$$

$$\frac{2}{3} G(P, Q) < \lg \frac{1}{r} < 2 G(P, Q),$$

falls P und Q den Kreis τ nicht verlassen. Dann folgt, wenn P auf s_1 liegt, wegen $u(P) = 1$ aus (5, 8)

$$(6, 8) \quad \frac{2}{3} < \int_{\bar{\omega}} \lg \frac{1}{r_{PQ}} d\mu(Q) < 2.$$

Daraus folgt nebenbei die Beschränktheit von γ und λ für hinreichend kleine Mengen bei beliebiger Wahl der Normierungsbelegung μ . Nun hat bekanntlich das Potential $u^*(P)$ die doppelte Eigenschaft, gleichzeitig die kleinste obere Grenze und die größte untere Grenze der Potentiale für alle möglichen Verteilungen der Gesamtmasse $\mu^*(\bar{\omega})$ auf $\bar{\omega}$ zu sein. Ist also P auf $\bar{\omega}$, so gilt

$$\lambda \cdot 1 = \int_{\bar{\omega}} \lg \frac{1}{r_{PQ}} d(\lambda \cdot \mu^*(Q)) \leq \text{Max} \int_{\bar{\omega}} \lg \frac{1}{r_{PQ}} d\mu(Q) < 2,$$

andererseits ist

$$\frac{2}{3} < \text{Min} \int_{\bar{\omega}} \lg \frac{1}{r_{PQ}} d\mu(Q) \leq \int_{\bar{\omega}} \lg \frac{1}{r_{PQ}} d(\lambda \mu^*(Q)) = \lambda,$$

also gilt schließlich

$$(7, 8) \quad \frac{2}{3} < \lambda < 2$$

für jede mögliche Wahl der Belegung $\mu(\omega)$ des normierten Potentials, oder

$$(8, 8) \quad \frac{2}{3} \gamma^* < \gamma < 2 \gamma^*.$$

Es ist nun für das folgende zweckmäßig, auch in der Potentialtheorie die Kapazität so zu definieren wie beim Operator L , d. h. nicht mittels der Funktion $\lg \frac{1}{r}$, welche zur üblichen Definition γ^* führt, sondern mittels der Greenschen Funktion $\Gamma(P, Q)$ der Laplaceschen Gleichung. Nur müssen wir den Nachweis erbringen, daß $\Gamma(P, Q)$ die Forderungen VII und VIII erfüllt. VII ist ja bekanntlich erfüllt. Wir müssen nun zeigen, daß sich das normierte Potential in der Form

$$u(P) = \int_{\bar{\omega}} \Gamma(P, Q) d\mu(Q)$$

darstellen läßt. $\bar{\omega}$ liegt im Innern des Kreises τ , der seinerseits konzentrisch zu κ ist. Dann ist offenbar aus Symmetriegründen, wenn s_r die Peripherie von τ bedeutet, die Funktion

$$v(P) = \int_{s_r} \Gamma(P, Q) ds$$

bis auf einen konstanten Faktor das normierte Potential von τ . $v(P)$ ist innerhalb τ konstant. Nun brauchen wir bloß die Massenverteilung auf s_r von der konstanten Dichte 1 auf den Rand von $\bar{\omega}$ auszufegen, was in genau der gleichen Weise erfolgt wie beim logarithmischen Linienpotential. Auf den

inneren Rand von $\bar{\omega}$ entfällt hierbei bekanntlich keine Belegung. Das neue Potential stimmt dann auf $\bar{\omega}$ mit v überein, hat also dort einen konstanten Wert. Die so entstehende Belegungsdichte φ ist, da $\bar{\omega}$ von analytischen Kurven berandet wird, eine stetige Funktion der Bogenlänge, das normierte Potential $\bar{u}(P)$ hat also die Form

$$\bar{u}(P) = \int \Gamma(P, Q) \varphi(s_Q) ds.$$

Für einen Randpunkt p von $\bar{\omega}$ gilt dabei wie beim logarithmischen Potential die Relation

$$-\frac{\partial \bar{u}}{\partial n} = \varphi(p) \quad (n = \text{innere Normale von } \omega^*).$$

Dies ist aber auf s_1 positiv. Damit ist die Eigenschaft VIII für Δu bewiesen. Wir haben also einen speziellen Fall eines Operators L vor uns, nämlich das Poissonsche Integral, und es gilt demnach für die sich so ergebende Kapazität

$$\bar{\gamma} = \int_{s_1} \varphi ds$$

die Beziehung (8, 8)

$$(9, 8) \quad \frac{1}{2} \gamma^* < \bar{\gamma} < 2 \gamma^*.$$

Aus (8, 8) und (9, 8) folgt leicht

$$(10, 8) \quad \frac{1}{2} \gamma < \gamma < 3 \bar{\gamma}.$$

Wir wollen dies nun für den Fall beliebiger abgeschlossener Punktmengen $\bar{\omega}$ in κ verallgemeinern. ω_n seien wieder die in § 6 zur Approximation von $\omega_0 - \omega^* = \kappa - \omega^*$ verwendeten, von analytischen Kurven begrenzten Punktmengen. Es war $\omega_n \supset \omega_{n+1} \supset \kappa - \omega^*$. γ_n sei die Kapazität von ω_n . Wegen (8, 8) sind die γ_n gleichmäßig beschränkt; denn daß die entsprechenden γ_n^* monoton abnehmen, ist bekannt. Es ist also

$$\gamma_n < 2 \gamma_1^*.$$

Faßt man die $\mu(\omega)$ auf als Belegungen der festen Menge ω_1 , so kann man wegen der gleichmäßigen Beschränktheit der γ_n eine schwach konvergente Teilfolge $\mu_{n_k}(\omega)$ auswählen. Wir erhalten so eine nichtnegative Grenzbelegung $\mu(\omega)$, deren Gesamtmasse nicht größer ist als $2 \gamma_1^*$. Die Grenzbelegung $\mu(\omega)$ erzeugt das normierte Potential $u(P)$ von $\bar{\omega}$. Dies folgt unmittelbar aus der schwachen Konvergenz, da von einem Index n_0 an P sicher nicht mehr in ω_n liegt, $G(P, Q)$ also in Q stetig auf ω_n ($n > n_0$) ist. Aus dem Umstande, daß die Kapazitäten γ_n^* für Δu gegen die Kapazität γ^* von $\bar{\omega}$ konvergieren und aus den Ungleichungen

$$\frac{1}{2} \gamma_n^* < \gamma_n < 2 \gamma_n^*$$

folgt wegen der schwachen Konvergenz

$$(11, 8) \quad \frac{1}{3} \gamma^* \leq \gamma \leq 2 \gamma^*$$

und ebenso

$$(12, 8) \quad \frac{1}{3} \bar{\gamma} \leq \gamma \leq 3 \bar{\gamma},$$

wo alle Kapazitäten sich jetzt auf die beliebige abgeschlossene Punktmenge $\bar{\omega}$ beziehen. Es sei noch bemerkt, daß die Ungleichungen (11, 8) und (12, 8) immer gelten, wie auch die Kapazität γ , die ja möglicherweise mit einer großen Willkürlichkeit behaftet sein kann, bestimmt sein mag.

§ 9.

Abschluß der allgemeinen Theorie.

Der Satz über die regulären Randpunkte.

Wir müssen nun das Wiensche Kriterium für die Regularität eines Randpunktes beweisen. Es lautet genau wie in meiner früheren Arbeit¹⁰⁾:

Satz 5. *Der Randpunkt p ist regulär oder nicht, je nachdem die Reihe*

$$(1, 9) \quad \sum_{n=n_0}^{\infty} n \gamma_n$$

divergiert oder konvergiert.

Die Kapazitäten γ_n sind folgendermaßen definiert: Sei λ ein positiver echter Bruch und $\kappa_{\lambda n}$ der abgeschlossene Kreis mit dem Radius λ^n um den Randpunkt p der gegebenen offenen Punktmenge Ω . Setzen wir

$$\Omega_n = \kappa_{\lambda n} - \Omega \kappa_{\lambda n}, \quad \omega_n = \Omega_n - \Omega_{n+1},$$

so sei

$$(2, 9) \quad u_n(P) = \int_{\omega_n} G(P, Q) d\mu_n(Q)$$

das normierte Potential von $\bar{\omega}_n$ für einen Kreis κ_0 um p aus $\bar{\Omega}$ und

$$\mu_n(\bar{\omega}_n) = \gamma_n$$

die Kapazität von $\bar{\omega}_n$. Ist n_0 hinreichend groß gewählt, so liegen alle Mengen $\bar{\omega}_n$ im Innern von κ_0 . Wegen VII, 2) gibt es eine Konstante M derart, daß für alle Punkte Q im Kreise $\kappa_{\lambda n_0}$ die Ungleichung

$$(3, 9) \quad \left| G(P, Q) - \lg \frac{1}{r_{PQ}} \right| = |w(P, Q)| < M \quad (P \in \kappa_0)$$

¹⁰⁾ Math. Zeitschr. 39 (1935), S. 532–559, im folgenden kurz mit R zitiert. Beim Beweis des Wienschen Kriteriums ist die Bezeichnung wie folgt abgeändert worden: statt T, K, E_n, e_n, E_a, v_n früher ist jetzt $\Omega, \kappa, \Omega_n, \bar{\omega}_n, \Omega_a, u_n$ geschrieben worden.

erfüllt ist. Ist ε eine der Ungleichung

$$0 < \varepsilon < 1$$

unterworfen, sonst beliebige Zahl, so wählen wir den Index m so groß, daß

$$(4, 9) \quad \lambda^{2m} < \frac{1}{2}, \quad M < \varepsilon \lg \frac{1}{r_{PQ}} \quad (P, Q \in \kappa_{2m})$$

ist. Dann folgt aus (3, 9)

$$(5, 9) \quad (1 - \varepsilon) \lg \frac{1}{r} < G(P, Q) < (1 + \varepsilon) \lg \frac{1}{r}.$$

Ist jetzt $n \geq 2m$, so folgt aus der Darstellung (2, 9), da $\mu_n(\omega)$ positiv monoton ist, und aus (5, 9)

$$(6, 9) \quad (1 - \varepsilon) \gamma_n \lg \frac{1}{\text{Max } r_{PQ}} \leq u_n(P) \leq (1 + \varepsilon) \gamma_n \lg \frac{1}{\text{Min } r_{PQ}} \quad (P \in \kappa_{2m}, Q \in \bar{\omega}_n).$$

Nun sei κ_α ein beliebiger abgeschlossener Kreis vom Radius α um p im Innern von κ_0 , $\Omega_\alpha = \kappa_\alpha - \Omega \kappa_\alpha$, $u_\alpha(P)$ das normierte Potential von Ω_α für κ_0 .

Um zu zeigen, daß das Kriterium dieses Paragraphen hinreichend ist, genügt es, nach Satz 4 zu zeigen, daß das normierte Potential $u_\alpha(P)$ die Beziehung

$$\lim_{P \rightarrow p} u(P) = 1$$

erfüllt, falls die Reihe (1, 9) divergiert. Wie in R nehmen wir an, daß die Reihe $\sum_i \gamma_i$ divergiert. Der Index m sei so groß, daß (4, 9) erfüllt ist, und außerdem κ_{2m} in κ_α enthalten ist. Es wird also (4, 9) verschärft in

$$(4, 9) a \quad \lambda^{2m} < \text{Min} \left(\frac{1}{2}, \alpha \right).$$

Da die beiden fundamentalen Abschätzungen (5, 9) und (6, 9) bewiesen sind, kann der weitere Beweis wörtlich aus meiner früheren Arbeit übernommen werden, und zwar von der Ungleichung (13) auf S. 543 an bis zur Mitte von S. 545.

Der Beweis der Umkehrung in R enthält eine Ungenauigkeit, nämlich die Abschätzung (21). Wir wollen ihn daher kurz unter Richtigstellung des betreffenden Schlusses wiederholen.

Sei $\sum_i \gamma_i$ konvergent. Wegen Bedingung VI ist $L_{1, \kappa_0}(P)$ im Mittelpunkt p von κ_0 positiv

$$L_{1, \kappa_0}(p) = \eta > 0.$$

Sei nun m so groß, daß wegen der Konvergenz von $\sum_i \gamma_i$

$$(7, 9) \quad \sum_{i=m}^{\infty} \gamma_i < \frac{1}{\lg \frac{1}{\lambda}} \cdot \frac{\eta}{8}$$

bleibt. Es wird also in der Abschätzung (18) von R auf der rechten Seite nur der Faktor η angehängt. Sei v_m das normierte Potential von $\Omega_m = \kappa_m - \Omega \kappa_m$, während u_m wie früher das normierte Potential von $\bar{\omega}_m = \bar{\Omega}_m - \bar{\Omega}_{m+1}$ bedeutet. Besitzt v_m in p den Grenzwert 1, so sei s_x der Umfang eines Kreises κ um p , auf welchem einschließlich s_x

$$v_m(P) > \frac{3}{4} \quad (P \in (\kappa + s_x) \Omega)$$

gilt. Jetzt sei $m'(>m)$ so groß, daß

$$\kappa_{2m'} \subset \kappa, v_{m'}(P) < \frac{1}{4} \quad (P \in s_x)$$

ist. $v_{m,m'}(P)$ sei das normierte Potential von $\Omega_m - \Omega_{m'}$. Es ergibt sich wieder wie in R

$$v_m \leq v_{m'} + v_{m,m'}.$$

Auf $s_x \Omega$ ist dann

$$\frac{3}{4} < v_m < \frac{1}{4} + v_{m,m'},$$

$$v_{m,m'}(P) > \frac{1}{2} \quad (P \in s_x \Omega).$$

Es ist also

$$2 v_{m,m'}(P) > 1 = L_{1,x}(P) > L_{1,x_0}(P)$$

auf $s_x \Omega$. Mit Hilfe der Oberfunktionen von $v_{m,m'}(P)$ folgert man leicht, daß auch in ganz $\kappa - \kappa(\Omega_m - \Omega_{m'})$

$$2 v_{m,m'}(P) \geq L_{1,x_0}(P)$$

im besonderen also

$$(8, 9) \quad v_{m,m'}(p) \geq \frac{1}{2} \eta$$

ist. Andererseits gilt in p

$$(9, 9) \quad v_{m,m'}(p) \leq \sum_{i=m}^{m'} u_i \leq (1 + \varepsilon) \sum_{i=m}^{m'} \gamma_i \lg \frac{1}{\lambda^{i+1}} = (1 + \varepsilon) \sum_{i=m}^{m'} i \gamma_i \lg \frac{1}{\lambda^{1+\frac{1}{i}}} \\ \leq 2 \sum_{i=m}^{\infty} i \gamma_i \lg \frac{1}{\lambda^2} < \frac{1}{2} \eta$$

nach (7, 9). (8, 9) und (9, 9) ergeben aber einen Widerspruch. $v_m(P)$ ist also für kein m stetig in p , p also kein regulärer Punkt. W. z. b. w.

Aus der Ungleichung (12, 8) folgt nun sofort, daß die Reihen

$$\sum_i i \gamma_i, \quad \sum_i i \bar{\gamma}_i$$

gleichzeitig divergieren bzw. konvergieren. Darans folgt endlich der

Satz 6. Die regulären Punkte bezüglich des Operators L sind identisch mit den regulären Punkten bezüglich Δu .

3. Anwendung auf elliptische Differentialgleichungen.

§ 10.

Die Bedingungen I bis VII.

Wir definieren jetzt den Operator L durch die Lösungen der in I betrachteten Funktionalgleichung

$$(1, 10) \quad L(u) = - \int_{\Omega} \frac{\partial u}{\partial n} ds + \int_{\omega} \left(\frac{\partial u}{\partial x} dA + \frac{\partial u}{\partial y} dB + u dC \right) = 0.$$

Die Funktionen $A(\omega)$, $B(\omega)$, $C(\omega)$ seien in einem beschränkten Gebiet Ω_0 definiert und genügen dort der Hölderbedingung (H) (vgl. I, S. 697)

$$\|F(\tau)\| < C \cdot r^{1+\alpha} \quad (\tau \text{ ein Kreis vom Radius } r)$$

mit dem für ganz Ω_0 geltenden Exponenten α . Die Konstante C der Bedingung soll für jeden in Ω_0 liegenden abgeschlossenen Bereich, aber nicht notwendig für ganz Ω_0 fest wählbar sein. Außerdem setzen wir noch voraus, daß in Ω_0 die Bedingung $(C_{a,b})$ (vgl. S. 714, I)

$$C(\omega) \leq 0, \quad \int_{\omega} \|dC\| > h \int_{\omega} (\|dA\| + \|dB\|)$$

erfüllt ist. Für die Konstante $\frac{1}{h}$ gilt dasselbe wie für die Hölderkonstante C . Das System \mathfrak{S} ist zunächst das System aller Kreise in Ω_0 , für welches die sukzessive Approximation gemäß Satz 1 (S. 698, I) durchführbar ist. Wir werden das System allerdings gleich noch etwas einschränken müssen. Offenbar sind die Bedingungen I, II, III und wegen $(C_{a,b})$ auch IV erfüllt. Die gleichgradige Stetigkeit (Bedingung V) folgt sofort aus der Abschätzung der ersten Ableitungen Du der Lösung u [vgl. Gleichung (76), S. 702 in I]

$$|Du| \leq \frac{H \cdot F}{R - \varrho}.$$

Dabei ist F eine obere Schranke für den Absolutbetrag der Randfunktion $f(s)$. Liegen die Funktionen $f(s)$ also unter einer festen Schranke F , so sind die ersten Ableitungen von u in einem festen Punkte gleichmäßig beschränkt, die Funktion selbst mithin gleichgradig stetig.

Auch Bedingung VI läßt sich leicht erledigen. Der erste Teil folgt unmittelbar aus Satz 3 von I (S. 710). Es ist noch zu zeigen, daß $L_{1,x}(P) > 0$ ist. Führen wir die sukzessive Approximation mit der Ausgangsfunktion $u_0 \equiv 1$ durch, so fallen alle Schwierigkeiten mit den Randableitungen fort und man kann den Index n_0 von Gleichung (8), S. 701 in I gleich 0 setzen, d. h. es gilt die Abschätzung

$$|u_0|, |Du_0| \leq 1,$$

$$|u - u_0| \leq \sum_{n=1}^{\infty} U_n < \frac{M}{1-M}.$$

Wählen wir $M < \frac{1}{2}$, so ist

$$|u - u_0| < 1$$

und

$$u > 0.$$

Ist also q eine Zahl zwischen Null und Eins, so bestimmen wir nun das System \mathfrak{S} endgültig so, daß wir von den früheren Kreisen nur die beibehalten, für welche die Konstante M kleiner als $\frac{q}{2}$ ausfällt. Ist die Hölderkonstante C für ganz Ω_0 fest wählbar, so sind dies alle Kreise aus Ω_0 , deren Radius unter einer festen hinreichend kleinen Zahl liegt.

Die Existenz einer Grundlösung (Bedingung VII) ist ebenfalls nach I gewährleistet.

§ 11.

Ein Eindeutigkeitssatz.

Es ist noch der Nachweis der Bedingung VIII zu erbringen, d. h. die Integraldarstellung des normierten Potentials zu beweisen. Die folgenden Hilfssätze stellen eine Verallgemeinerung und zum Teil eine Vereinfachung entsprechender Sätze von R dar, welche in der Befreiung von der Existenz zweiter Ableitungen begründet ist. Die Beweise lassen sich zum Teil übertragen, sind aber vollständig durchgeführt. Wir beweisen zunächst folgende Verallgemeinerung des Hilfssatzes 3 von R:

Hilfssatz 1. ω^{*11} sei ein beschränktes Gebiet, dessen Rand s aus endlich vielen einfachen geschlossenen analytischen Kurven als Komponenten zusammengesetzt ist. s_0 sei der Außenrand, s_1 die Gesamtheit der übrigen Komponenten. Die Funktion $w(P)$ sei stetig in $\omega^* + s$ und in ω^* eine reguläre Lösung von (1, 10). Außerdem seien ihre ersten Ableitungen in $\omega^* + s_1$ stetig, und es gelte

$$w = 0 \text{ auf } s_0, \quad \frac{\partial w}{\partial n} = 0 \text{ auf } s_1.$$

Dann ist

$$w(P) \equiv 0.$$

Nehmen wir an, daß w in einem Punkte p_0 von s_1 ein positives Maximum $w(p_0) = m$ besitzt. Wegen Bedingung $(C_{a,b})$ ist in ω^*

$$(1, 11)$$

$$w(P) < m.$$

¹¹⁾ Die Bezeichnung ist in Übereinstimmung mit der Bezeichnung von § 6 gewählt.

Wie früher bilden wir eine Umgebung U von p_0 durch eine Funktion $z' = \varphi(z)$ ($z' = x' + iy'$) schlicht und konform so ab, daß das (analytische) Randstück von s_1 in U in ein Stück der y' -Achse übergeht. Die Normalableitung von w geht wieder in eine solche über und verschwindet also auf $x' = 0$. Außerdem wird unter Benützung der Konformität der Abbildung

$$L(w) = L'(w) = - \int_{s'} \frac{\partial w}{\partial \bar{n}'} ds' + \int_{\omega'} \left(\frac{\partial u}{\partial x'} dA' + \frac{\partial u}{\partial y'} dB' + u dC' \right),$$

$$A' = \int_{\omega} \frac{\partial x'}{\partial x} dA + \frac{\partial x'}{\partial y} dB, \quad B' = \int_{\omega} \frac{\partial y'}{\partial x} dA + \frac{\partial y'}{\partial y} dB, \quad C' = C \quad (\omega' \text{ Bild von } \omega).$$

Ähnlich wie im früheren Falle setzen wir sämtliche Funktionen im Bilde U' von U über die y' -Achse hinaus fort durch die Vorschrift

$$(2, 11) \quad w(x', y') = w(-x', y'), \quad A'(\bar{\omega}') = -A'(\omega'), \quad B'(\bar{\omega}') = B'(\omega'), \\ C'(\bar{\omega}') = C(\omega'),$$

wo $\bar{\omega}'$ das Spiegelbild der Menge ω' an der y' -Achse ist. Es sei nun κ ein Kreis um $p_0(z_0)$, der samt seinem Rande im Regularitätsgebiet der Abbildungsfunktion $z' = \varphi(z)$ liegt, κ' sei sein Bild. Dann gibt es, wie man z. B. mit Hilfe des Schwarzschen Lemmas einsieht, eine positive Konstante M derart, daß für jeden Kreis τ' in κ' vom Radius r'

$$r < Mr'$$

gilt, wo r der Radius des kleinsten Kreises τ ist, welcher das Urbild von τ' enthält.

Die Funktionen A' , B' , C' genügen dann in κ' einer Hölderbedingung. Z. B. ist

$$\|A'(\tau')\| < 2CNr^{1+\alpha} < 2CNM^{1+\alpha}r'^{1+\alpha},$$

wenn N eine obere Schranke für die Ableitung $\left| \frac{d\varphi}{dz} \right|$ in κ ist.

Setzen wir nun

$$(3, 11) \quad W(z') = w_0(z') + \frac{1}{2\pi} \int_{\tau'} \Gamma(z', \zeta') \left[\frac{\partial w}{\partial x'} dA' + \frac{\partial w}{\partial y'} dB' + w dC' \right],$$

wobei τ' ein Kreis in κ' mit dem Mittelpunkt z'_0 ist, und die Potentialfunktion w_0 dieselben Randwerte wie w besitzt. $\Gamma(z', \zeta')$ ist die Greensche Funktion des Kreises τ' . $W(z')$ ist offenbar symmetrisch zur y' -Achse und besitzt, wie in I bewiesen, stetige erste Ableitungen. Daher ist

$$(4, 11) \quad \frac{\partial W}{\partial x'} = 0 \quad (x' = 0).$$

Wir behaupten, daß für jede Menge ω' , die von einer rektifizierbaren Kurve s' begrenzt wird, die Gleichung gilt

$$(5, 11) \quad \int_{s'} \frac{\partial w}{\partial n} ds = \int_{\omega'} \left(\frac{\partial w}{\partial x} dA' + \frac{\partial w}{\partial y} dB' + w dC' \right) = \int_{s'} \frac{\partial W}{\partial n} ds.$$

(5, 11) gilt sicher, wenn ω' ganz in einer der beiden Hälften des Kreises κ' liegt. Denn für w gilt sie, weil w reguläre Lösung nach Voraussetzung ist, und für W , weil es die Summe einer Potentialfunktion und eines logarithmischen Potentials ist. Ragt nun ω' in beide Hälften hinein, so wird es durch die y' -Achse in zwei Teile ω'_1 und ω'_2 zerlegt, für welche (5, 11) erfüllt ist. Da, wie schon bemerkt, auf der y' -Achse $\frac{\partial w}{\partial n}$, und nach (4, 11) auch $\frac{\partial W}{\partial n}$ verschwindet, so ergibt sich durch Zusammensetzen der entsprechenden Formeln für ω'_1 und ω'_2 schließlich die behauptete Gleichung (5, 11). Es ist also

$$\int_{s'} \frac{\partial (W - w)}{\partial n} ds = 0$$

für alle einfachen, geschlossenen, streckbaren Kurven s' in τ' . $W - w$ ist also harmonisch, und, da es auf dem Rande von τ' verschwindet, auch im Innern von τ' gleich Null:

$$W \equiv w.$$

w ist also Lösung der Integralgleichung

$$w = w_0 + \frac{1}{2\pi} \int_{\tau'} \Gamma(z', \zeta') \left(\frac{\partial w}{\partial x'} dA' + \frac{\partial w}{\partial y'} dB' + w dC' \right)$$

und mithin reguläre Lösung von $L'(w) = 0$ in τ' . Nach dem Maximumprinzip mußte also in τ'

$$w = m$$

sein im Widerspruch zu (1, 11).

§ 12.

Zwei weitere Hilfssätze ¹²⁾.

ω^* , s_0 , s_1 mögen dieselbe Bedeutung haben wie im vorigen Paragraphen. Nur sei jetzt der Durchmesser von ω^* hinreichend klein, so daß Satz 4 von I anwendbar ist. ω_0 sei das Innere von s_0 , ω das Innere von s_1 , also $\omega + s_1 + \omega^* = \bar{\omega} + \omega^* = \omega_0$. $G(P, Q)$ sei die Greensche Funktion für ω_0 .

¹²⁾ Vgl. hierzu S. 534 in R.

Hilfssatz 2. Für jede beschränkte integrable Funktion φ der Bogenlänge t auf s_1 ist

$$(1, 12) \quad W(s) = \int_{s_1} \frac{\partial G}{\partial n_s}(P_s, Q_t) \varphi(t) dt$$

eine hölderstetige Funktion der Bogenlänge s mit festem Exponenten.

Beweis. Es bedeute $\lg \frac{1}{r} + w(P, Q)$ die bei der Konstruktion von $G(P, Q)$ benutzte Grundlösung; die Funktion u in der Zerlegung

$$(2, 12) \quad G = \lg \frac{1}{r} + w + u$$

ist mithin eine reguläre Lösung in ω_0 . Entsprechend dieser Zerlegung zerlegen wir auch $W(s)$

$$W = W_1 + W_2 + W_3.$$

Durch $s' \neq s''$ sind zwei Punkte P', P'' irgendeiner Komponente von s_1 bestimmt. Sei

$$|s' - s''| = \delta, \quad r' = \overline{Q_t P'}, \quad r'' = \overline{Q_t P''}.$$

Dann sei σ der Teil von s_1 , auf welchem

$$\text{Min}(r', r'') \leq \delta^{\frac{\alpha}{3}};$$

auf $s_1 - \sigma$ ist dann

$$(3, 12) \quad \text{Min}(r', r'') \geq \delta^{\frac{\alpha}{3}}.$$

σ besteht aus allen Kurvenstücken von s_1 , die in der Vereinigung zweier Kreise um die Punkte s' und s'' mit dem Radius $\delta^{\frac{\alpha}{3}}$ enthalten sind. Bezüglich $W_3(s)$ ist nichts zu beweisen, da u in ω_0 regulär ist, und die ersten Ableitungen von u nach I gleichmäßig in bezug auf s_1 einer H-Bedingung genügen. Die Differenz

$$|W_1(s') - W_1(s'')|$$

ist bereits in R (S. 557) abgeschätzt worden. Danach ist

$$(4, 12) \quad |W_1(s') - W_1(s'')| \leq C_- \delta^{\frac{1}{4}}.$$

Bleibt noch

$$|W_2(s') - W_2(s'')|$$

abzuschätzen. Setzen wir

$$(5, 12) \quad W_2(s'') - W_2(s') = V_\sigma + V_{s_1 - \sigma}$$

entsprechend der Zerlegung von s_1 . Gemäß (13 b), S. 713, I gilt die Abschätzung

$$|Dw| < \frac{C}{r^{1-\alpha}}.$$

Mithin wird

$$(6, 12) \quad |V_\sigma| \leq \int_a^\sigma \left(\left| \frac{\partial w}{\partial n_{\sigma'}} \right| + \left| \frac{\partial w}{\partial n_{\sigma''}} \right| \right) |\varphi| dt \leq C \int_a^\sigma \left(\frac{1}{r'^{1-\alpha}} + \frac{1}{r''^{1-\alpha}} \right) dt.$$

Es genügt etwa $\int_a^\sigma \frac{dt}{r'^{1-\alpha}}$ abzuschätzen. σ liegt nach der Bemerkung zu (3, 12) sicher in einem Kreise mit dem Radius $\delta + \delta^{\frac{\alpha}{3}}$ um P' , für hinreichend kleines δ also in einem Kreise τ_ρ vom Radius $\rho = 2\delta^{\frac{\alpha}{3}}$ um P' . Legen wir den Koordinatenursprung in den Punkt P' , und nehmen die Tangente an σ in P' als x -Achse, so ist s_1 in einer gewissen Umgebung dieses Punktes darstellbar in der Form

$$(7, 12) \quad y = a_0 x^2 + a_1 x^3 + \dots \quad (|x| \leq \eta < 0).$$

Wir wählen die Konstante η , die nur von der Kurve s_1 abhängt, noch so, daß für eine beliebige positive Zahl m

$$|y'| < m < 0 \quad (|x| \leq \eta)$$

gilt. Ist δ hinreichend klein, so enthält τ_ρ nur Punkte des durch (7, 12) dargestellten Stückes von s_1 , das natürlich aus einem Kurvenzug besteht. Aus (7, 12) folgen für r und die Bogenlänge t leicht die Darstellungen

$$r = \sqrt{x^2 + y^2} = x \sqrt{1 + \left(\frac{y}{x}\right)^2} = x(1 + O(x^2)),$$

$$t = \int_0^x \sqrt{1 + y'^2} dx = x(1 + O(x^2)),$$

$$r = t(1 + O(x^2)).$$

Es ist also in τ_ρ etwa

$$r > \frac{t}{2}, \quad \int_a^\sigma \frac{dt}{r'^{1-\alpha}} < \int_0^{\frac{2\rho}{t}} \frac{dt}{\left(\frac{t}{2}\right)^{1-\alpha}} = \frac{2\rho^\alpha}{\alpha} < \frac{4}{\alpha} \delta^{\frac{\alpha^2}{3}}.$$

Mithin ist

$$(8, 12) \quad |V_\sigma| = O(\delta^{\frac{\alpha^2}{3}}).$$

Nun die Abschätzung von

$$(9, 12) \quad V_{s_1-\sigma} = \int_{s_1-\sigma}^\sigma \left(\frac{\partial w}{\partial n_{\sigma'}} - \frac{\partial w}{\partial n_{\sigma''}} \right) \varphi(t) dt.$$

Da die Grundlösung $G^* = \lg \frac{1}{r} + w$ für $P \neq Q$ eine Lösung von $L(w) = 0$ ist, hat w die Form

$$w = \frac{1}{2\pi} \int_{\omega_0} \Gamma(P, R) dF(R, Q),$$

$$F(w, Q) = \int_{\omega} \left(\frac{\partial G^*(R, Q)}{\partial x_R} dA(R) + \frac{\partial G^*(R, Q)}{\partial y_R} dB(R) + G^*(R, Q) dC(R) \right).$$

Gehen wir nun an die Abschätzung der Differenz

$$(10, 12) \quad \frac{\partial w}{\partial n_{s''}} - \frac{\partial w}{\partial n_{s'}} = \frac{1}{2\pi} \int_{\omega_0} \left(\frac{\partial \Gamma(P'', R)}{\partial n_{s''}} - \frac{\partial \Gamma(P', R)}{\partial n_{s'}} \right) dF(R, Q),$$

wobei $Q = Q_t$ im Integral $V_{s_1-\sigma}$ die Punkte von $s_1 - \sigma$ zu durchlaufen hat. Wir schätzen zunächst die Differenz

$$\left| \frac{\partial \Gamma}{\partial n_{s''}} - \frac{\partial \Gamma}{\partial n_{s'}} \right|$$

ab. P', P'' liegen auf s_1 , haben also einen positiven Abstand von s_0 . Setzen wir nun

$$\Gamma = \lg \frac{1}{r} + v,$$

so genügt die Potentialfunktion v gleichmäßig in R einer Abschätzung

$$(11, 12) \quad \left| \frac{\partial v}{\partial n_{s''}}(P'', R) - \frac{\partial v}{\partial n_{s'}}(P', R) \right| < C \cdot \delta \quad (\delta = |s'' - s'|).$$

Weiterhin ist, $\overline{RP'} = \varrho'$, $\overline{RP''} = \varrho''$ gesetzt,

$$\frac{\partial \lg \varrho''}{\partial n_{s''}} - \frac{\partial \lg \varrho'}{\partial n_{s'}} = \cos(n_{s''}, \varrho'') \left(\frac{1}{\varrho''} - \frac{1}{\varrho'} \right) + \frac{\cos(n_{s''}, \varrho'') - \cos(n_{s'}, \varrho')}{\varrho'}.$$

Die erste Differenz ist offenbar absolut kleiner als

$$\frac{\delta}{\varrho' \varrho''}.$$

Die zweite ergibt sich aus dem Mittelwertsatz bei Berücksichtigung der Beschränktheit der Krümmung als ein $\frac{O(\delta)}{\varrho' \varrho''}$. Zusammen mit (11, 12) folgt also

$$\left| \frac{\partial \Gamma}{\partial n_{s''}} - \frac{\partial \Gamma}{\partial n_{s'}} \right| \leq C \frac{\delta}{\varrho' \varrho''}.$$

Nun ist wegen (13b) in I

$$|DG^*| \leq \frac{C}{\varrho} \quad (\varrho = \overline{RQ}).$$

Mithin erhalten wir für die Differenz (10, 12)

$$\left| \frac{\partial w}{\partial n_{s''}} - \frac{\partial w}{\partial n_{s'}} \right| \leq C \cdot \delta \int_{\omega_0} \frac{\|dA\| + \|dB\| + \|dC\|}{\varrho' \varrho'' \varrho} = C \delta \int_{\omega_0} \frac{dH(R)}{\varrho \varrho' \varrho''}.$$

q, q', q'' sind die Verbindungsgeraden des Punktes R mit Q, P', P'' . R variiert in ω_0, Q auf $s_1 - \sigma$. Es ist

$$|q' - q''| \leq |s' - s''| = \delta,$$

τ sei ein Kreis mit dem Radius $\varrho^* = \frac{1}{2} \delta^{\frac{\alpha}{3}}$ um Q , dann ist

$$\varrho \geq \frac{1}{2} \delta^{\frac{\alpha}{3}} \quad (R \in \omega_0 - \tau),$$

$$\varrho', \varrho'' \geq \frac{1}{2} \delta^{\frac{\alpha}{3}} \quad (R \in \tau).$$

Folglich ist unter Benutzung einer Abschätzung aus I (S. 712)

$$\delta \int_{\omega_0 - \tau} \frac{dH}{\varrho \varrho' \varrho''} \leq \frac{2\delta}{\delta^{\frac{\alpha}{3}}} \int_{\omega_0} \frac{dH}{\varrho' \varrho''} < C \frac{\delta}{\delta^{\frac{\alpha}{3}}} \cdot \delta^{\alpha-1} = C \delta^{\frac{2\alpha}{3}},$$

$$\delta \int_{\tau} \frac{dH}{\varrho \varrho' \varrho''} \leq \frac{4\delta}{\delta^{\frac{2\alpha}{3}}} \int_{\tau} \frac{dH}{\varrho} \leq C \delta^{\frac{\alpha}{3}} \varrho^{*\alpha} = C \delta^{\frac{\alpha}{3}(1+\alpha)} < C \delta^{\frac{\alpha}{3}},$$

$$\left| \frac{\partial w}{\partial n_{s''}} - \frac{\partial w}{\partial n_{s'}} \right| \leq C \delta \int_{\omega_0} \frac{dH}{\varrho \varrho' \varrho''} < C \delta^{\frac{\alpha}{3}},$$

$$(12, 12) \quad |V_{s_1 - \sigma}| \leq C \delta^{\frac{\alpha}{3}}.$$

(4, 12), (8, 12), (12, 12) und die Bemerkung über $W_3(s)$ liefern schließlich die im Hilfssatz 2 behauptete H -Stetigkeit des Integrals (1, 12).

Hilfssatz 3. Die Belegungsfunktion $\varphi(t)$ sei jetzt gleichmäßig hölderstetig auf s_1 , d. h. $|\varphi(t') - \varphi(t'')| < C|t' - t''|^\gamma$ ($\gamma > 0$) mit festem C und γ . Schließen wir s_1 in eine einfache geschlossene im Innern von ω^* verlaufende Kurve S ein, so besitzt das Potential

$$v(P) = \int_{s_1} G(P, Q_t) \varphi(t) dt$$

in ω sowie in dem von S und s_1 begrenzten Gebiet gleichmäßig hölderstetige erste Ableitungen, die mithin auf den Rändern derselben Gebiete stetige Grenzwerte besitzen.

Für den Fall, daß $G(P, Q)$ durch die Funktion $\lg \frac{1}{r}$ ersetzt wird, ist die Aussage des Satzes bereits bewiesen (vgl. Lichtenstein, Enc. II C 3, S. 201/202). Zerlegen wir $G(P, Q)$ gemäß (2, 12) in

$$G = \lg \frac{1}{r} + w + u,$$

so ist der Beweis nur für die Funktion

$$\int_{s_1} w(P, Q_t) \varphi(t) dt$$

durchzuführen, da die reguläre Lösung u im Innern von S hölderstetige erste Ableitungen besitzt. Nach I ist wie vorhin, wenn wieder

$$\lg \frac{1}{r} + w = G^*$$

gesetzt wird,

$$w(P, Q) = \frac{1}{2\pi} \int_{\omega_0} \Gamma(P, R) \left[\frac{\partial G^*(R, Q)}{\partial x_R} dA(R) + \frac{\partial G^*(R, Q)}{\partial y_R} dB(R) + G^*(R, Q) dC(R) \right].$$

Wir führen den Beweis für die Funktion

$$w_1(P, Q) = \frac{1}{2\pi} \int_{\omega_0} \Gamma(P, R) \frac{\partial G^*(R, Q)}{\partial x_R} dA(R)$$

durch. Dann ist

$$\begin{aligned} (13, 12) \quad \int_{s_1} w_1(P, Q_t) \varphi(t) dt &= \frac{1}{2\pi} \int_{s_1} \varphi(t) dt \int_{\omega_0} \Gamma(P, R) \frac{\partial G^*(R, Q_t)}{\partial x_R} dA(R) \\ &= \frac{1}{2\pi} \int_{\omega_0} \Gamma(P, R) dA(R) \int_{s_1} \frac{\partial G^*(R, Q_t)}{\partial x_R} \varphi(t) dt. \end{aligned}$$

Die Erlaubtheit der Vertauschung folgt wegen der Beschränktheit von $\varphi(t)$ leicht aus dem Fubini-Tonellischen Satze, da das erste Integral auch bei absolut genommenem Integranden existiert, was für das innere Integral in I (S. 711 unten) bewiesen wurde. Nun ist das Integral

$$\int_{s_1} \frac{\partial G^*(R, Q_t)}{\partial x_R} \varphi(t) dt = \Phi(R)$$

als Funktion von R beschränkt. Dies folgt für den Anteil $\frac{\partial}{\partial x_R} \int_{s_1} \lg \frac{1}{r} \varphi(t) dt$

nach Lichtenstein (l. c., S. 201) aus der Hölderstetigkeit von $\varphi(t)$, für den Anteil $\int_{s_1} \frac{\partial w(R, Q_t)}{\partial x_R} \varphi(t) dt$ ebenfalls unter Benutzung der Hölderstetigkeit von $\varphi(t)$ aus der Abschätzung (13b) I

$$|Dw| < \frac{N}{r^{1-\alpha}}.$$

Setzen wir

$$\int_{\omega_0} \Phi(R) dA(R) = A_1(\omega),$$

so genügt $A_1(\omega)$ wegen der Beschränktheit von Φ derselben H -Bedingung wie $A(\omega)$, und (13, 12) schreibt sich

$$\int_{s_1} w_1(P, Q_t) \varphi(t) dt = \frac{1}{2\pi} \int_{\omega_0} \Gamma(P, R) dA_1(R).$$

Diese Funktion besitzt aber nach I in ω_0 hölderstetige erste Ableitungen.

Damit wäre der Beweis erbracht. Aus ihm folgt zugleich die Gültigkeit der bekannten Randrelation für die Normalableitung des Potentials der einfachen Belegung, falls diese einer H -Bedingung genügt. Denn wir haben soeben gezeigt, daß die ersten Ableitungen des Anteils

$$\int_{s_1} (w + u) \varphi(t) dt$$

stetig bleiben beim Durchgange durch s_1 . $v(P)$ verhält sich daselbst also genau

wie das logarithmische Potential $\int_{s_1} \lg \frac{1}{r} \varphi(t) dt$. D. h. es gilt

$$(14, 12) \quad \frac{\partial v}{\partial n_s} = -\pi \varphi(s) + \int_{s_1} \frac{\partial G(P_s, Q_t)}{\partial n_s} \varphi(t) dt,$$

wo n_s die ins Äußere von ω gerichtete Normale bedeutet. Die entsprechende Beziehung gilt aber ebenso in bezug auf die Gebiete ω . Bedeuten also $\frac{\partial v}{\partial n_i}$ und $\frac{\partial v}{\partial n_o}$ die Grenzwerte der ins Äußere von ω weisenden Normalableitungen auf s_1 bezüglich der Gebiete ω und ω^* , so folgt aus (14, 12) und der entsprechenden Gleichung für ω die Sprungrelation

$$(15, 12) \quad \frac{\partial v}{\partial n_i} - \frac{\partial v}{\partial n_o} = 2\pi \varphi.$$

§ 13.

Abschluß des Beweises.

Wir sind nun in der Lage, den angekündigten Satz über die Darstellung normierter Potentiale und damit das Erfülltsein sämtlicher an unseren speziellen Operator L gestellten Forderungen zu beweisen.

Satz 7. *Das Gebiet ω^* mit dem Rande $s = s_0 + s_1$ genüge den Bedingungen von § 12. $u(P)$ sei die nach Satz 1 von I existierende Funktion mit den Randwerten 0 auf s_0 und 1 auf s_1 . Dann gibt es eine und nur eine stetige nicht-negative Belegungsfunktion $\varphi(t)$, mittels deren sich $u(P)$ durch das folgende Integral darstellen läßt*

$$(1, 13) \quad u(P) = \int_{s_1} G(P, Q_t) \varphi(t) dt.$$

Beweis. Wegen (14, 12) muß $\varphi(t)$ eine Lösung der Integralgleichung

$$(2, 13) \quad \frac{\partial u}{\partial n_s} = -\pi \varphi(s) + \int_{s_1} \frac{\partial G}{\partial n_s}(P_s, Q_t) \varphi(t) dt$$

sein. Nach Satz 1 von I ist $\frac{\partial u}{\partial n_s}$ noch auf dem Rande hölderstetig. Der Kern $\frac{\partial G}{\partial n_s}$ ist wegen der Stetigkeit von $\frac{\cos(n_s, r)}{r}$ auf s_1 und wegen (13b) in I

nur uneigentlich singular und es gilt die Fredholmsche Alternative. Besitzt die homogene Gleichung

$$-\pi \varphi(s) + \int_{s_1} \frac{\partial G}{\partial n_s}(P, Q_t) \varphi(t) dt = 0$$

eine nichttriviale stetige Lösung $\bar{\varphi}(t)$, so ist nach dem Hilfssatz 2 von § 12 $\bar{\varphi}(s)$ hölderstetig. Für das Potential

$$v(P) = \int_{s_1} G(P, Q_t) \bar{\varphi}(t) dt$$

gilt nach der Schlußbemerkung des vorigen Paragraphen die Gleichung (14, 12).

Die Normalableitung $\frac{\partial v}{\partial n_s}$ auf s_1 , die nach Hilfssatz 3 ja existieren muß, ist also gleich Null und daher nach dem Eindeutigkeitssatz auch v selbst. Daraus wiederum und aus der Stetigkeit von v in ganz ω_0 folgt wegen des Maximumprinzips, daß v auch in ω verschwindet, und daher auch die Ableitung $\frac{\partial v}{\partial n_i}$. Nach (15, 12) ist demnach

$$\bar{\varphi}(t) \equiv 0.$$

die homogene Integralgleichung also nur trivial lösbar. Folglich hat die inhomogene Gleichung (2, 13) eine und nur eine stetige Lösung $\varphi(t)$. Diese ist nach Hilfssatz 2 sogar hölderstetig, da, wie schon erwähnt, die Ableitung $\frac{\partial u}{\partial n_s}$ in (2, 13) ebenfalls hölderstetig ist. Also hat das mit dieser Belegung gebildete Potential

$$v(P) = \int_{s_1} G(P, Q_t) \varphi(t) dt$$

nach Hilfssatz 3 noch auf s_1 eine stetige Normalableitung, die wegen (2, 13) mit $\frac{\partial u}{\partial n_s}$ übereinstimmt. Da v ebenso wie u auf s_0 verschwindet, so sind u und v nach dem Eindeutigkeitssatz überhaupt identisch. Es ist noch zu zeigen, daß die Belegung $\varphi(t)$ nichtnegativ ist. Dies folgt aber aus der Sprungrelation (15, 12), wenn man bedenkt, daß wegen der Gültigkeit des Maximumprinzips v sowohl in ω , wie in ω^* höchstens gleich 1 sein kann; die Ableitungen $\frac{\partial v}{\partial n_i}$, $-\frac{\partial v}{\partial n_s}$, und damit auch $\varphi(t)$ sind also nichtnegativ, da die Normalen nach ω^* hinein gerichtet sind.

(Eingegangen am 1. 5. 1942.)

Zur „wahrscheinlichkeitstheoretischen Stabilisierung“ beim Erneuerungsproblem.

Von

Hans Schwarz in Göttingen.

Das Erneuerungsproblem ist neuerdings von Herrn H. Richter¹⁾ einer eingehenden mathematischen Behandlung unterzogen worden. Zur weiteren Informierung über diesen Fragenkreis, sowie über seine Bedeutung in der Bevölkerungstheorie sei deshalb auf jene Arbeit und das dortige umfassende Literaturverzeichnis verwiesen.

Herr H. Richter hat einen neuen Stabilisierungsbegriff eingeführt. Er spricht von „wahrscheinlichkeitstheoretischer Stabilisierung“, wenn der Erwartungswert der bis zu einem Zeitpunkt zu erneuernden Elemente asymptotisch proportional der Zeit und der mittlere Fehler von kleinerer Ordnung als die Zeit wächst. Es wird bei kontinuierlicher Betrachtungsweise gezeigt, daß wahrscheinlichkeitstheoretische Stabilisierung stattfindet, wenn die Ausscheidefunktion quadratintegrierbar ist und ihr zweites Potenzmoment existiert.

In der vorliegenden Arbeit soll gezeigt werden, daß dies schon der Fall ist, wenn nur das erste Potenzmoment der Ausscheidefunktion, die sog. mittlere Verweilzeit, existiert. Der Beweis wird zunächst bei der diskontinuierlichen Betrachtungsweise mit Hilfe eines Satzes Tauberscher Art über Potenzreihen geführt und dann ganz analog bei der kontinuierlichen Betrachtung mit Hilfe einer Übertragung jenes Satzes auf Laplace-Transformationen. Weiter werden bei der diskontinuierlichen Betrachtung der oben genannte Erwartungswert und der mittlere Fehler mit Hilfe der Additionssätze für Erwartungswerte und mittlere Fehlerquadrate berechnet, da es interessiert, wie diese in der Wahrscheinlichkeitstheorie übliche Schlußweise im vorliegenden Fall angewandt werden kann.

Betrachten wir eine Gesamtheit von Elementen, die einer diskontinuierlichen oder kontinuierlichen Ausscheidordnung unterliegen: p_n bzw. $p(t) dt$ sei die Wahrscheinlichkeit, daß ein Element n bzw. t Zeiteinheiten nach Eintritt in die Gesamtheit ausscheidet (Ausscheidfunktion). Im Zeitpunkt Null soll sich die Gesamtheit aus lauter neuen Elementen zusammensetzen.

¹⁾ H. Richter, Untersuchungen zum Erneuerungsproblem. Math. Annalen 118, Heft 2.

Ist dann φ_n bzw. $\varphi(t)dt$ die Wahrscheinlichkeit, daß bei konstant gehaltenem Bestand im Zeitpunkt n bzw. t ein Element ausscheidet, also erneuert werden muß (Erneuerungsfunktion), so bestehen bekanntlich zwischen p und φ die Beziehungen, bei diskontinuierlicher Betrachtung:

$$(1) \quad \varphi_n - p_n = \sum_{r=1}^{n-1} p_{n-r} \varphi_r, \quad \varphi_r = \sum_{n=r}^{\infty} \varphi_{n-r} p_n, \quad \varphi_1 = p_1,$$

bei kontinuierlicher Betrachtung:

$$(2) \quad \varphi(t) - p(t) = \int_0^t p(t-\tau) \varphi(\tau) d\tau = \int_0^t \varphi(t-\tau) p(\tau) d\tau.$$

Die Funktionen p sind Wahrscheinlichkeitsverteilungen mit dem Argumentbereich $(0, \infty)$. Es ist also vorauszusetzen:

$$(3) \quad p_n \geq 0, \quad \mu_0 = \sum_{n=1}^{\infty} p_n = 1,$$

$$(4) \quad p(t) \geq 0, \quad \mu_0 = \int_0^{\infty} p(t) dt = 1.$$

Weiter soll noch $p(t)$ in jedem endlichen Intervall beschränkt sein, also

$$(5) \quad p(t) \leq M \text{ in } 0 \leq t \leq t_0^2).$$

Wir wenden uns nun dem diskontinuierlichen Fall zu und wollen zunächst den Erwartungswert Φ_n und den mittleren Fehler σ_n der in dem Zeitintervall $(0, n)$ ausscheidenden Elemente berechnen. Für wahrscheinlichkeitstheoretische Betrachtungen ist es zweckmäßig, sich den konstanten Bestand nur aus einem Element bestehend zu denken, das nach gewissen, der Ausscheidungsfunktion unterliegenden Intervallen ersetzt wird.

φ_i ist die Wahrscheinlichkeit, daß im Zeitpunkt i ein Element ausscheidet, $1 - \varphi_i$ die Wahrscheinlichkeit, daß dies nicht der Fall ist. Also ist

$$(6) \quad 1 \cdot \varphi_i + 0 \cdot (1 - \varphi_i) = \varphi_i$$

die zu erwartende Anzahl der im Zeitpunkt i ausscheidenden Elemente, und nach dem Additionssatz für Erwartungswerte folgt

$$\Phi_n = \sum_{i=1}^n \varphi_i.$$

Der Additionssatz für mittlere Fehlerquadrate lautet hier:

$$\sigma_n^2 = \sum_{i=1}^n s_i^2 + \sum_{i \neq k} r_{ik} s_i s_k,$$

²⁾ Vgl. H. Richter, a. a. O.

wobei s_i das mittlere Fehlerquadrat bezüglich des Zeitpunktes i ist und r_{ik} der Korrelationskoeffizient zwischen dem Ausscheiden in i und k . Zunächst findet man ohne weiteres

$$s_i^2 = (1 - \varphi_i)^2 \varphi_i + (0 - \varphi_i)^2 (1 - \varphi_i) = \varphi_i (1 - \varphi_i).$$

Zur Berechnung von $r_{ik} s_i s_k$ ist die Kenntnis der Wahrscheinlichkeiten erforderlich, daß sowohl in i als auch in k ein Element ausscheidet bzw. daß in i , nicht aber in k ein Element ausscheidet usw. Sei zunächst $i < k$, dann seien die vier Wahrscheinlichkeiten in unmißverständlicher Weise mit w_{11} , bzw. w_{10} , w_{01} , w_{00} bezeichnet. Es ist

$$\begin{aligned} w_{11} &= \varphi_i \varphi_{k-i}, \\ w_{10} &= \varphi_i (1 - \varphi_{k-i}), \\ w_{01} &= \varphi_k - \varphi_i \varphi_{k-i}, \\ w_{00} &= 1 - \varphi_i - \varphi_k + \varphi_i \varphi_{k-i}. \end{aligned}$$

Die ersten beiden Gleichungen sind ohne weiteres einzusehen, die letzten beiden ergeben sich etwa aus

$$\begin{aligned} w_{11} + w_{01} &= \varphi_k, \\ w_{10} + w_{00} &= 1 - \varphi_k. \end{aligned}$$

Also ist für $i < k$

$$\begin{aligned} r_{ik} s_i s_k &= (1 - \varphi_i) (1 - \varphi_k) w_{11} + (1 - \varphi_i) (0 - \varphi_k) w_{10} + \\ &\quad + (0 - \varphi_i) (1 - \varphi_k) w_{01} + (0 - \varphi_i) (0 - \varphi_k) w_{00} \\ &= \varphi_i (\varphi_{k-i} - \varphi_k). \end{aligned}$$

Ist $k < i$, so braucht man nur die Indizes zu vertauschen und erhält

$$r_{ik} s_i s_k = \varphi_k (\varphi_{i-k} - \varphi_i).$$

und nach dem Additionssatz ergibt sich nun

$$\begin{aligned} \sigma_n^2 &= \sum_{i=1}^n \varphi_i (1 - \varphi_i) + \sum_{i < k} \varphi_i (\varphi_{k-i} - \varphi_k) + \sum_{k < i} \varphi_k (\varphi_{i-k} - \varphi_i) \\ &= \sum_{i=1}^n \varphi_i - \sum_{i=1}^n \varphi_i^2 - \sum_{i \neq k} \varphi_i \varphi_k + 2 \sum_{i < k} \varphi_i \varphi_{k-i}, \end{aligned}$$

$$(7) \quad \sigma_n^2 = 2 \sum_{r=1}^{n-1} \Phi_r \varphi_{n-r} + \Phi_n - \Phi_n^2.$$

Wahrscheinlichkeitstheoretische Stabilisierung findet also statt, wenn

$$(8) \quad \frac{1}{n} \Phi_n \rightarrow \text{konst.} \quad \text{und} \quad \frac{1}{n} \sigma_n \rightarrow 0 \quad \text{für} \quad n \rightarrow \infty,$$

und es soll untersucht werden, unter welchen Voraussetzungen über die Ausscheidefunktion p_n dies der Fall ist. Es wird sich herausstellen, daß die Existenz der mittleren Verweilzeit

$$(9) \quad \mu_1 = \sum_{r=1}^{\infty} r p_r$$

hinreicht.

Zu diesem Zweck betrachten wir die Lösung der Differenzengleichung (1) mittels Potenzreihen³⁾. Bildet man mit der komplexen Variablen z die Potenzreihe

$$(10) \quad g(z) = \sum_{r=1}^{\infty} p_r z^r,$$

so ist $g(z)$ wegen (3) im Innern des Einheitskreises analytisch und auf dem Kreise stetig, und es ist $g(1) = 1$. Im Innern des Einheitskreises hat $1 - g(z)$ keine Nullstellen, denn für $|z| < 1$ ist

$$|g(z)| \leq \sum_{r=1}^{\infty} p_r |z|^r < 1.$$

Die Funktion

$$(11) \quad f(z) = \frac{g(z)}{1 - g(z)}$$

ist also im Innern des Einheitskreises regulär. Sie läßt sich um $z = 0$ in eine Potenzreihe entwickeln

$$(12) \quad f(z) = \sum_{r=1}^{\infty} \varphi_r z^r,$$

die in $|z| < 1$ konvergiert. Aus $f(z) - g(z) = f(z)g(z)$ folgt nun

$$\sum_{n=1}^{\infty} (\varphi_n - p_n) z^n = \sum_{n=2}^{\infty} \sum_{r=1}^{n-1} p_{n-r} \varphi_r z^n,$$

also ist $\varphi_n = p_n$ die Lösung von (1).

Es sei nun

$$(13) \quad P_n = \sum_{r=n}^{\infty} p_r, \quad G(z) = \sum_{r=1}^{\infty} P_r z^r,$$

dann ist unter der Voraussetzung (9) wieder $G(z)$ analytisch in $|z| < 1$ und stetig auf $|z| = 1$, und es ist

$$(14) \quad \mu_1 = \sum_{r=1}^{\infty} P_r, \quad G(1) = \mu_1.$$

³⁾ Vgl. H. Hadwiger, Untersuchungen über das asymptotische Verhalten rekurrenter Zahlenreihen. Mite. d. Vereinig. schweiz. Versicherungsmathematiker 34 (1937).

Aus der Identität

$$p, z^r = \frac{z-1}{z} P, z^r + P, z^{r-1} - P, z^{r+1} z^r$$

folgt durch Summation über r von 1 bis ∞

$$(15) \quad 1 - g(z) = \frac{1-z}{z} G(z)$$

in $|z| \leq 1$ (Abelsche partielle Summation). Hieraus folgt in Verbindung mit (11) und (12)

$$(1-z) \sum_{r=1}^{\infty} \varphi, z^r = \frac{zg(z)}{G(z)}$$

und wegen (14)

$$(16) \quad \lim_{z \rightarrow 1} (1-z) \sum_{r=1}^{\infty} \varphi, z^r = \frac{1}{\mu_1}.$$

Wir benutzen nun den folgenden Satz Tauberscher Art:

Sei $a_n \geq 0$, $\sum_{r=1}^{\infty} a, z^r$ in $|z| < 1$ konvergent, und für ein $\alpha \geq 0$ strebe

$$(1-z)^{\alpha} \sum_{r=1}^{\infty} a, z^r \rightarrow A$$

für $z \rightarrow 1$ auf der reellen Achse, so gilt für $n \rightarrow \infty$

$$\frac{1}{n^{\alpha}} \sum_{r=1}^n a, \rightarrow \frac{A}{\Gamma(\alpha+1)} \quad ^4).$$

Ist a_n eine monoton wachsende Folge, so gilt sogar

$$\frac{1}{n^{\alpha-1}} a_n \rightarrow \frac{A \alpha}{\Gamma(\alpha+1)} \quad ^5).$$

(Den Zusatz werden wir später benötigen.)

Da $\varphi_n \geq 0$ ist, wie man aus (1) ohne weiteres ersieht, ergibt sich aus (16)

$$(17) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \Phi_n = \frac{1}{\mu_1},$$

und damit ist der erste Teil von (8) bewiesen.

Um zu zeigen, daß auch $\lim_{n \rightarrow \infty} \frac{1}{n} \sigma_n = 0$, betrachten wir

$$\psi_n = \sum_{r=1}^{n-1} \Phi, \varphi_{n-r} = \sum_{r=1}^{n-1} \varphi, \Phi_{n-r}.$$

Zunächst ist ψ_n monoton wachsend, denn sei etwa $m < n$, so ist

$$\psi_n - \psi_m = \sum_{r=1}^{m-1} \varphi, (\Phi_{n-r} - \Phi_{m-r}) + \sum_{r=m}^{n-1} \varphi, \Phi_{n-r} > 0,$$

⁴⁾ Hardy u. Littlewood, Lond. M. S. Proc. (2) 13 (1914), S. 174.

⁵⁾ G. Doetsch, Math. Zeitschr. 11 (1921), S. 161.

da Φ_n monoton nie abnimmt. In $|z| < 1$ ist nun

$$F(z) = \sum_{r=1}^{\infty} \Phi_r z^r = \frac{f(z)}{1-z}$$

und

$$h(z) = \sum_{r=1}^{\infty} \varphi_r z^r = F(z) f(z)$$

oder

$$(1-z)^3 h(z) = [(1-z) f(z)]^2,$$

und wegen (16) folgt

$$(18) \quad \lim_{z \rightarrow 1} (1-z)^3 \sum_{r=1}^{\infty} \varphi_r z^r = \frac{1}{\mu_1^2}.$$

Da φ_n positiv und monoton wachsend ist, kann man hier nach dem Zusatz des obigen Satzes von Hardy und Littlewood schließen:

$$(19) \quad \lim_{n \rightarrow \infty} \frac{1}{n^2} \varphi_n = \frac{1}{2\mu_1^2}.$$

Aus (17) und (19) folgt endlich

$$\frac{1}{n^2} \sigma_n^2 = \frac{2}{n^3} \varphi_n + \frac{\Phi_n}{n^3} - \left(\frac{\Phi_n}{n} \right)^2 \rightarrow 0 \quad \text{für } n \rightarrow \infty,$$

also

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sigma_n = 0,$$

und damit ist auch der zweite Teil von (8) bewiesen.

Für die Integralgleichung (2) soll jetzt noch der entsprechende Satz bewiesen werden. Unter den zu Beginn gemachten Voraussetzungen über die Ausscheidefunktion (4) und (5) gilt:

Wenn die mittlere Verweilzeit

$$(20) \quad \mu_1 = \int_0^{\infty} t p(t) dt$$

existiert, so findet wahrscheinlichkeitstheoretische Stabilisierung statt, d. h. es gilt dann

$$(21) \quad \frac{1}{t} \Phi(t) \rightarrow \frac{1}{\mu_1} \quad \text{und} \quad \frac{1}{t} \sigma(t) \rightarrow 0 \quad \text{für } t \rightarrow \infty,$$

wobei

$$(22) \quad \Phi(t) = \int_0^t \varphi(\tau) d\tau,$$

$$(23) \quad \sigma^2(t) = 2 \int_0^t \Phi(\tau) \varphi(t-\tau) d\tau + \Phi(t) - \Phi^2(t)$$

der Erwartungswert und das mittlere Fehlerquadrat der im Zeitintervall $(0, t)$ ausscheidenden Elemente sind. Herr H. Richter hat in der erwähnten Arbeit diese Werte durch Aufstellen der Verteilung und mit Hilfe von erzeugenden Funktionen berechnet.

Der Beweis verläuft mit Hilfe von Laplace-Transformationen ganz analog. Es ist

$$(24) \quad g(z) = \int_0^{\infty} p(t) e^{-zt} dt$$

wegen (4) eine in $R(z) > 0$ ⁶⁾ analytische und in $R(z) \geq 0$ stetige Funktion und $g(1) = 1$. Ist $\varphi(t)$ die immer vorhandene eindeutige Lösung von (2), so ist

$$(25) \quad f(z) = \int_0^{\infty} \varphi(t) e^{-zt} dt$$

konvergent in $R(z) > 0$, und es gilt dort

$$(26) \quad f(z) = \frac{g(z)}{1 - g(z)} \quad 7).$$

In $R(z) > 0$ ergibt sich nun für die Transformierte $F(z)$ von $\Phi(t)$:

$$F(z) = \int_0^{\infty} \Phi(t) e^{-zt} dt = \frac{f(z)}{z} \quad 8).$$

und für die Transformierte $h(z)$ von $\psi(t) = \int_0^t \Phi(\tau) \varphi(t - \tau) d\tau$ nach dem Faltungssatz ⁹⁾

$$(27) \quad h(z) = \int_0^{\infty} \psi(t) e^{-zt} dt = F(z) f(z),$$

also

$$(28) \quad z^2 h(z) = [zf(z)]^2.$$

Unter der Voraussetzung (20), daß μ_1 existiert, ist

$$(29) \quad G(z) = \int_0^{\infty} P(t) e^{-zt} dt \quad \text{mit} \quad P(t) = \int_t^{\infty} p(t) dt$$

⁶⁾ Realteil von z .

⁷⁾ Vgl. H. Richter, a. a. O.

⁸⁾ Vgl. G. Doetsch, Theorie und Anwendung der Laplacetransformation. Berlin 1937. S. 148.

⁹⁾ Anm. 8, S. 161.

wieder in $R(z) > 0$ analytisch, in $R(z) \geq 0$ stetig und

$$(30) \quad G(1) = \mu_1 = \int_0^{\infty} P(t) dt.$$

Durch partielle Integration findet man hier in $R(z) \geq 0$

$$(31) \quad g(z) = 1 + z G(z).$$

Daher ist nun nach (26), (28) und (31)

$$z f(z) = \frac{g(z)}{G(z)},$$

$$z^3 h(z) = \left[\frac{g(z)}{G(z)} \right]^3$$

in $R(z) > 0$ und wegen (30)

$$(32) \quad \lim_{z \rightarrow 0} z f(z) = \frac{1}{\mu_1},$$

$$(33) \quad \lim_{z \rightarrow 0} z^3 h(z) = \frac{1}{\mu_1^3}.$$

Um jetzt aus diesem Verhalten der Transformierten bei $z = 0$ auf das Verhalten von $\varphi(t)$ bzw. $\psi(t)$ für $t \rightarrow \infty$ zu schließen, benutzen wir folgende Übertragung des obigen Satzes von Hardy und Littlewood auf Laplace-Transformationen¹⁰⁾:

Sei $a(t) \geq 0$, $\int_0^{\infty} a(t) e^{-zt} dt$ in $R(z) > 0$ konvergent, und für $\alpha \geq 0$ strebe

$$z^{\alpha} \int_0^{\infty} a(t) e^{-zt} dt \rightarrow A$$

für $z \rightarrow 0$ auf der reellen Achse, so gilt für $t \rightarrow \infty$

$$\frac{1}{t^{\alpha}} \int_0^t a(t) dt \rightarrow \frac{A}{\Gamma(\alpha+1)}.$$

Ist außerdem $a(t_1) \leq a(t_2)$ für $0 \leq t_1 \leq t_2$, so gilt sogar

$$\frac{1}{t^{\alpha-1}} a(t) \rightarrow \frac{A \alpha}{\Gamma(\alpha+1)}.$$

Daß $\varphi(t) \geq 0$ ist, erkennt man etwa aus der Darstellung von $\varphi(t)$ durch eine Neumannsche Reihe. Daher ist $\Phi(t)$ positiv und monoton nie abnehmend,

¹⁰⁾ Ann. 8, S. 208.

und genau wie früher folgt dies auch für $\psi(t)$. Also können wir aus (32) und (33) schließen, daß

$$\lim_{t \rightarrow \infty} \frac{1}{t} \Phi(t) = \frac{1}{\mu_1} \text{ }^{11)},$$

$$\lim_{t \rightarrow \infty} \frac{1}{t^2} \Psi(t) = \frac{1}{\mu_1^2},$$

und es gilt weiter

$$\frac{1}{t^2} \sigma^2(t) = \frac{2 \Psi(t)}{t^2} + \frac{\Phi(t)}{t^2} - \left(\frac{\Phi(t)}{t} \right)^2 \rightarrow 0 \text{ für } t \rightarrow \infty,$$

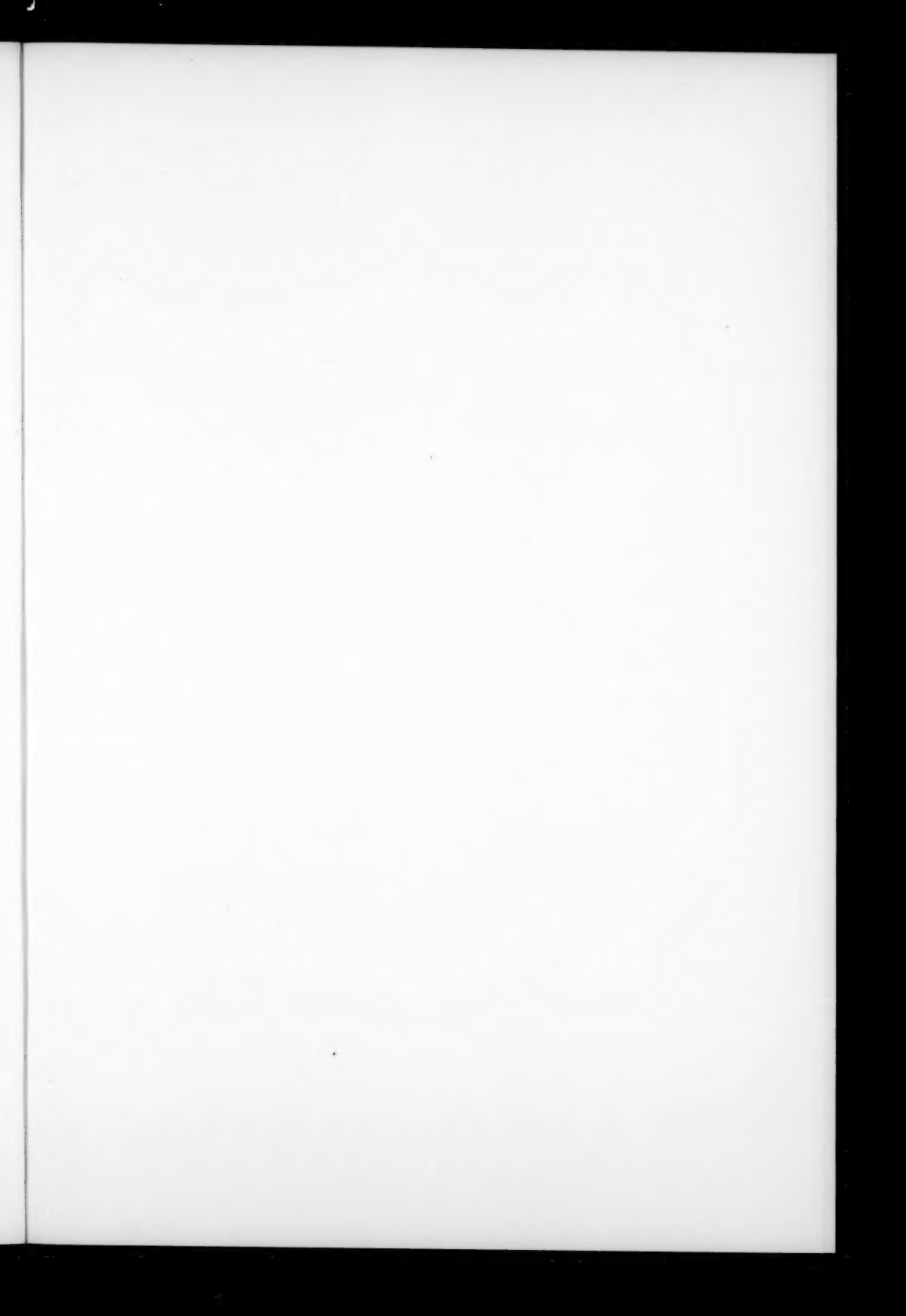
oder

$$\lim_{t \rightarrow \infty} \frac{1}{t} \sigma(t) = 0,$$

womit auch für die kontinuierliche Betrachtungsweise erwiesen ist, daß wahrscheinlichkeitstheoretische Stabilisierung stattfindet, wenn nur die mittlere Verweilzeit existiert.

¹¹⁾ Anm. während der Korrektur: In einem Referat, Zentralblatt f. Math. u. Grenzgeb., Bd. 26, Seite 127, weist H. Hadwiger auf die Möglichkeit dieses Beweises hin.

(Eingegangen am 8. 6. 1942.)





C15

0	E	E	E	□
1	3	2	8	
2	6	3	2	
3	2	3	5	
4	5	3	8	
5	5	5	5	
6	3	5	8	



0	E	E	E	□
1	3	2	8	
2	6	3	2	
3	2	3	5	
4	5	3	8	
5	5	5	5	
6	3	5	8	

